

# **Causation and Modern Philosophy**

**Edited by Keith Allen  
and Tom Stoneham**



# **Causation and Modern Philosophy**



# Routledge Advances in the History of Philosophy

## 1. Concepts and Reality in the History of Philosophy

Tracing a Philosophical Error from Locke to Bradley  
Fiona Ellis

## 2. Autobiography as Philosophy

The Philosophical Uses of Self-Presentation  
Edited by Thomas Mathien and D. G. Wright

## 3. Causation and Modern Philosophy

Edited by Keith Allen and Tom Stoneham



# **Causation and Modern Philosophy**

**Edited by Keith Allen  
and Tom Stoneham**



**Routledge**

Taylor & Francis Group

New York London



First published 2011  
by Routledge  
270 Madison Ave, New York, NY 10016

Simultaneously published in the UK  
by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

© 2011 Taylor & Francis

The right of the editors to be identified as the authors of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

Typeset in Sabon by IBT Global.  
Printed and bound in the United States of America on acid-free paper by IBT Global.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

*Library of Congress Cataloging in Publication Data*

A catalog record has been requested for this book.

ISBN13: 978-0-415-88355-9 (hbk)  
ISBN13: 978-0-203-83011-6 (ebk)



# Contents

<i>Acknowledgements</i>	vii
<b>Introduction</b> KEITH ALLEN AND TOM STONEHAM	1
<b>1 Galileo: Reflections on Failure</b> DAVID WOOTTON	13
<b>2 Primary and Secondary Causes in Descartes' Physics</b> TAD M. SCHMALTZ	31
<b>3 Causation and the Cartesian Reduction of Motion: God's Role in Grinding the Gears</b> WILLIAM EATON AND ROBERT HIGGERSON	48
<b>4 Spinoza's Conatus as an Essence-Preserving, Attribute-Neutral Immanent Cause: Toward a New Interpretation of Attributes and Modes</b> ERIC SCHLIESSER	65
<b>5 Are Mind-Body Relations Natural and Intelligible? Some Early Modern Perspectives</b> PAULINE PHEMISTER	87
<b>6 Hobbes's Redefinition of the Commonwealth</b> TIMOTHY STANTON	104
<b>7 Hume, Causal Realism, and Free Will</b> PETER MILLICAN	123

vi	<i>Contents</i>	
8	<b>Pouring New Wine into Old Skin: The Meaning of Hume's Necessary Connexions</b> CONSTANTINE SANDIS	166
9	<b>Is Causation a Relation?</b> BORIS HENNIG	188
10	<b>Kant on Causal Knowledge: Causality, Mechanism, and Reflective Judgment</b> ANGELA BREITENBACH	201
11	<b>Regularities All the Way Down: Thomas Brown's Philosophy of Causation</b> STATHIS PSILLOS	220
12	<b>Causality and Causal Induction: The Necessitarian Theory of Lady Mary Shepherd</b> MARTHA BRANDT BOLTON	242
	<i>List of Contributors</i>	263
	<i>Index</i>	267



# Acknowledgements

This volume arose out of the British Society for the History of Philosophy's conference 'Causation: 1500-2000', hosted by the Department of Philosophy at the University of York in March 2008. We would like to thank the British Society for the History of Philosophy, Routledge for supporting the conference, and Mike Beaney and Richard Flockemann for their work organising it. We would also like to thank the team at Routledge, particularly Erica Wetter, Diana Castaldini, and Ryan Kenney, for their help producing the volume.



# Introduction

*Keith Allen and Tom Stoneham*

The chapters in this volume chart the development of views of causation in the modern period. The chapters span the period from Galileo, Descartes, and their Scholastic predecessors to the reactions of Thomas Brown and Lady Mary Shepherd to the ‘Leslie affair’, the controversy that arose in 1805 surrounding the appointment of John Leslie, a mathematician and proponent of Hume’s views of causation, to a chair at Edinburgh University. The chapters illustrate the central importance of causation in modern philosophy, situating debates about the nature of causation in relation to issues in natural science, metaphysics, epistemology, philosophy of mind, theology, and politics.

The chapters in the volume can be divided into three parts. The chapters in the first part consider the development of views of causation in the seventeenth century, exploring in particular the legacies of Aristotelian and Cartesian theories of causation. Hume’s theory of causation is the focus of the second part. Finally, the chapters in the third part investigate post-Humean developments in thinking about causation.

## I. THE SEVENTEENTH CENTURY

Causation became an important subject in the seventeenth century in light of developments in natural philosophy, and particularly the turn away from Aristotelian science, with its interest in final causes, towards mechanism, with its exclusive focus on efficient causes. If any one idea can be said to be essential to mechanistic explanation, it is the thought that the only causal interactions between physical objects occur when two bodies are in motion relative to each other and they touch. Correlatively, science becomes the project of discovering the laws of motion and impact. One of the leading figures at the beginning of this mechanical revolution was Galileo. In “Galileo: Reflections on Failure”, Wootton argues that Galileo in fact held contradictory views on causation which reflect the position that he occupies at this historical turning point. On the one hand, Galileo explains the



practical difficulties associated with producing vacuums in terms of the need to overcome an occult nonmechanical force, a *horror vacui*—even though the conceptual resources to explain these difficulties in mechanist terms were available to Galileo. Elsewhere, however, Galileo eschews nonmechanical explanations in favour of purely mechanistic causal explanations, which he should have known simply did not work, perhaps most notably where the truth of Copernicanism is at stake, as in the case of Galileo’s explanation of the tides in terms of the earth’s movement. Wootton argues that these two failures on Galileo’s part—the Copernican explanation of the tides and the appeal to a *horror vacui*—are evidence of his confused thinking about causation. On the occasions when Galileo focuses explicitly on the nature of causation, he suggests one of two views: he either thinks of causation in mechanistic terms, as when he famously draws a version of the distinction between primary and secondary qualities; or he thinks of causation in interventionist terms, whereby the relationship between cause and effect is such that if you take away the cause, you take away the effect. But typically, Galileo is not interested in causation as such. Wootton likens Galileo’s attitude towards causation to that of a driver’s attitude towards the windscreen of their car: ‘unless something goes wrong, you look through it not at it’ (p. 22).

Subsequent debates about causation in the early modern period are shaped to a large extent by the more systematic metaphysical and epistemological framework inherited from Descartes. Central to the Cartesian system are two fundamental distinctions: between infinite God and particular finite substances; and amongst particular finite substances, between material substance (*res extensa*) and thinking substance (*res cogitans*). Thinking substances have affinities to both God and material substance: insofar as man is made in God’s image, thinking substances share with God the essential attribute of thought; insofar as man is made, thinking substances share with material substances a dependence on God.

This tripartite distinction immediately raises questions about the relations between the different kinds of substance. Exactly what does the dependence of created substances on God consist in? And how do particular created substances relate to each other? Descartes famously argues in the Third Meditation that all finite substances depend on God for their creation and continued conservation, a distinction that Descartes argues is merely a distinction of reason, because ‘the same power and action are needed to preserve anything at each individual moment of its duration as would be required to create that thing anew if it were not yet in existence’.<sup>1</sup> Nevertheless, God’s creation and conservation of particular substances is consistent with three more specific views of God’s relation to his Creation, which ascribe to God different levels of involvement in the unfolding of events within the world: mere conservationism, concurrentism, and occasionalism.

In “Primary and Secondary Causes in Descartes’ Physics”, Schmaltz argues that Descartes is a *mere conservationist* about God’s contribution



to the production of natural effects. This is the view that substances are dependent on God as primary cause only for their creation and conservation, and these substances are themselves the immediate causes of changes in other substances.

Mere conservationism differs in this respect from occasionalism and concurrentism, the two competing views of God's relation to his Creation that were abroad in the late Scholastic and early modern periods. According to occasionalism, God alone is the complete, exclusive, and immediate cause of all changes in nature; 'secondary causes' provide merely the 'occasion' for God's action.<sup>2</sup> To say that God is the *complete* cause is to say that He is alone sufficient for the production of the effect. He is the *exclusive* cause of all natural changes insofar as nothing else is necessary for the production of these effects. Finally, the claim that God is the *immediate* cause of all natural changes leaves room for a derivative notion of 'cause', which allows for the literal truth of causal ascriptions to particular created substances. As such, occasionalists need not deny that particular substances can causally affect each other, so long as the truth of these mediate causal ascriptions depends upon the truth of immediate ascriptions of causal powers to God: in this sense, occasionalism is consistent with the truth of ascriptions of causal influence to particular bodies in the same way that the claim that we immediately perceive only ideas is often argued to be consistent with the claim that we nevertheless perceive, albeit mediately, ordinary material objects.

Concurrentism represents a *via media* between occasionalism and mere conservationism, according to which God 'concurs' in the natural production of change. This view tries to respect the occasionalist's ascription of all causal agency to God, whilst acknowledging that particular created substances are more than merely mediate, derivate causes, by allowing that (at least in certain circumstances) complete causes do not have to be exclusive: just as both the author and their pen can jointly cause the writing on the page, so God and body can jointly cause another body to move.

The mere conservationism that Schmaltz attributes to Descartes requires ascribing to particular substances causal powers that are individually sufficient to produce their effects. This might seem problematic at least in the case of material substances, given Descartes's identification of material substance with extension, which is often understood as an inert geometrical quality; to attribute force to Cartesian bodies might therefore seem to involve appealing to the kind of Scholastic 'occult qualities' that Descartes and his successors aimed to avoid. Schmaltz argues, however, that this would be a mistake. Central to Schmaltz's interpretation is the distinction between geometrical qualities as they exist as ideas in the understanding and enduring geometrical qualities as they exist in reality. According to Schmaltz, material substances are not secondary causes in virtue of having powers over and above their extension. Rather, the physically instantiated geometrical properties of matter themselves constitute the powers



to produce changes in other material objects, differing depending on the strength of their tendency to endure. This tendency to endure does not differ from the modes of extension that themselves endure, and is ultimately grounded in God, at least insofar as God conserves the total quantity of the strength to endure in the world.

In direct opposition to Schmaltz, Eaton and Higginson defend an occasionalist interpretation of Descartes. Eaton and Higginson argue that occasionalism was entailed by what they describe as ‘solid Cartesian principles’: principles that Descartes and later Cartesians would have been reluctant to give up, and which are supported by key primary sources. Moreover, they argue that Descartes was himself committed to occasionalism, at least regarding body-body causation. Eaton and Higginson cite textual evidence, drawing mainly on Descartes’s correspondence with More, of Descartes’s commitment to three claims that jointly entail occasionalism: that motion is a mode of body; that ‘genuine’ causal interaction requires the transfer of motion from one body to another; and that modes cannot be transferred from one substance to another.

The interpretative debate raises numerous questions, not least of which is the consistency of Descartes’s views across his lifetime: whereas the evidence that Eaton and Higginson cite in support of their claim that Descartes is an occasionalist comes mainly from the late 1640s, towards the end of Descartes’s career, Schmaltz’s interpretation focuses on the major published works from the early 1640s and draws on the much earlier *Le Monde*. But whatever Descartes’s actual views, more general questions about how causation *should* be thought of within a Cartesian framework remain.

Drawing on the work of the later Cartesian La Forge, for instance, Eaton and Higginson argue that the occasionalism, to which the solid Cartesian principles they identify lead, has important theoretical benefits. In particular, occasionalism provides a way of reducing local motion to other modes of extension. Given that God is required to explain the conservation of matter anyway, it is possible to regard motion as ‘a sort of divine teleportation in which an object is *re-created* at successive locations’ (p. 49). Eaton and Higginson argue that such a reduction of motion to location and duration exhibits the key theoretical virtues of parsimony and intelligibility. It is more parsimonious than Descartes’s conception, because it reduces the number of basic modes that need to be attributed to matter to explain their interactions. It is more intelligible than Descartes’s conception of motion, which itself replaces the unintelligible Scholastic conception of motion (‘Motion is the actuality of a potential being in so far as it is potential’<sup>3</sup>) by reductively explaining motion in terms of something (at least supposedly) more intelligible: the divine re-creation of matter.

Schmaltz and Eaton and Higginson focus on Cartesian treatments of body-body interactions. Accounting for mind-body interactions within a Cartesian framework raises further issues. Some of these questions are addressed in the chapters by Schliesser and Phemister.



An apparent lacuna in Descartes's account of God's role in the conservation of created substances provides the background to Schliesser's account of Spinoza's conception of *Conatus*. Central to Descartes's account of body-body interactions are his laws of motion, described in Part II of the *Principles of Philosophy*. These laws of motion seemingly presuppose internal principles that preserve the identity of substances over time. In the case of material substances, this is the principle of 'inertia'. But what is the analogue of 'inertia' for mental substances?

By a process of 'reverse-conceptual-engineering', Schliesser argues that Spinoza's *Conatus* is intended to fill this gap. According to Schliesser, *Conatus* is an 'attribute neutral, essence preserving, immanent cause'. It is *attribute neutral* in the sense that it is that which explains the identity of entities over time, irrespective of whether they are mental or material; as such, *Conatus* addresses the perceived lacuna in Descartes's account of thinking substance. It is *essence preserving* in the sense that it determines the identity over time of the *entity* which the constant *states* are states of. Finally, it is an *immanent cause* in that it expresses in a 'certain and determinate' way God's power.

In addressing problems that arise within the Cartesian framework, Spinoza ends up radically reforming this system, replacing Descartes's distinction between three types of substance by an all-encompassing monism. Spinoza effectively eliminates questions about the causal relations between God, bodies, and minds by arguing that there is just one substance, one that is differently modified in mental and material respects. A similar transformation of the Cartesian framework, though to somewhat different effect, is undertaken by Leibniz, as Phemister's contribution to the volume explains.

Phemister examines Leibniz's response to the problem of mind-body interaction. The view standardly ascribed to Descartes is that mind and body are distinct substances that interact causally: the body in producing sensations or perceptions in the mind, the mind in turn producing movement in the body. But the heterogeneity of mind, the essence of which is thought, and body, the essence of which is extension, puts pressure on the intelligibility of genuine causal interaction between minds and bodies.

Leibniz's response to the problem of mind-body interactionism appeals to a pre-established harmony between minds and bodies. Leibniz rejects both Cartesian interactionism and Malebranchean occasionalism on the grounds that they require causal interaction between heterogeneous substances. As such, both violate a requirement on the human intelligibility of divine action: interactionism because it requires causal communication between mind and body, and occasionalism because it requires causal interaction between God and created substances (but contrast Eaton and Higginson). Instead, according to Leibniz, changes within souls and bodies—in their sequences of perceptions and movements—are caused, respectively, by those very souls and bodies acting according to preordained general laws. On the one hand, 'intra-substance causation' satisfies a requirement on



intelligible causal interaction, that causes be ‘like’ (in some suitable respect) their effects. On the other hand, the pre-established harmony between causal processes in minds and bodies (noncausally) explains the intricate arrangement of bodies, which, as Damaris Masham (a friend of Locke and defender of interactionism) argues, might otherwise appear mysterious. According to Leibniz, an intelligible connection between perceptions of colour, pain, other sensible qualities, and motions in objects is secured by insensible (‘petites’) perceptions in the soul, which correspond to, or ‘mirror’, motion in the insensible parts of the object perceived.

As Phemister points out, whereas tensions within the Cartesian framework lead Spinoza to the position that there is one infinite substance which is simultaneously modified in respect of thought and extension, Leibniz argues for a pluralistic view in which there are an indefinite number of individual substances that both perceive and are subject to physical forces. Where Spinoza and Leibniz agree is that these different sequences of modifications are ultimately expressions of one underlying essence.

Although Galileo and Descartes set the agenda for many of the discussions of causation in the early modern period, their views did not arise *ex nihilo*. Whereas Wootton situates Galileo’s thinking about causation in relation to his commitment to Copernicanism and (general) hostility to Aristotelianism, and Schmaltz emphasises the extent to which Descartes’s view of God’s relation to nature was informed by discussions amongst late Scholastics, a different aspect of the Aristotelian legacy—Aristotle’s distinction between four kinds of cause—is discussed in more detail in Stanton’s contribution to the volume. Whereas many of the chapters in the volume reflect the importance of causation to the metaphysical underpinnings of the science of the early modern period, Stanton’s chapter highlights the role of differing conceptions of causation in politics and political philosophy, arguing that Hobbes’s theory of the Commonwealth is grounded in a metaphysical thesis about the nature of causation.

Hobbes’s theory of causation reconfigures the fourfold classification of causes found in Aristotle. Aristotle distinguished the *material* cause, or the matter out of which an effect comes; the *formal* cause, or that which produces the essence of the effect; the *efficient* cause, or the movement responsible for the effect; and the *final* cause, or the end or purpose for which something is brought into existence. Hobbes reduces this fourfold classification to a bipartite distinction, identifying the efficient cause with the agent (that which acts), the material cause with the patient (that which is acted upon), and assimilating final and formal causes to efficient causes. On the one hand, final causes are nothing more than desires of the agent that a particular type of thing exist, and therefore nothing over and above the total efficient cause. On the other hand, and like other early modern mechanists, Hobbes was sceptical of the existence of ‘forms’ or ‘essences’ (as distinct from the things themselves) that formal causes in Aristotelian philosophy appeal to. Instead, Hobbes identifies essences with definitions,



or words that signify the way in which we think about things, that at best cause knowledge of the things that they signify. In contrast to the Cartesians, Hobbes's rejection of Aristotelianism shows affinity with Hume's rejection of substance metaphysics and the famous 'fork' between matters of fact and relations of ideas.

Stanton argues that Hobbes's understanding of causation has important implications for understanding his account of the Commonwealth: a union of men who willingly submit themselves to the will of a sovereign. There is a long tradition, dating from the time of Hobbes's contemporary, the Royalist pamphleteer Robert Filmer, of seeing Hobbes as trying to offer an account of the way in which individuals generate a Commonwealth from the state of nature—as on Filmer's own account, according to which the sovereign's authority over his subjects is explained by the communication to him of an essence through his generative causal history, ultimately tracing back to Adam, the original husband, father, and king. But because Hobbes identifies essences with definitions, he rejects the possibility of transmitting essences. Instead of explaining the sovereign's authority in terms of the *generation* of the commonwealth, Stanton argues that the authority of the sovereign derives from the sovereign's role in protecting and *preserving* the lives of the subjects who comprise the commonwealth. Hence the causal explanation of the authority of the sovereign is reversed: it derives from the sovereign's role as cause rather than effect.

## II. HUME

Although Galileo and Descartes might originally have set the agenda for modern discussions of causation, it was radically reconfigured by Hume. But exactly what Hume's views are is controversial. The chapters by Millican and Sandis continue the lively debate between Old and New Humeans over the question of whether Hume was a sceptical realist about causation.

Millican laid down a 'clear challenge' to the New Humeans in 2007 to explain Hume's determinist argument that, because it meets the conditions set down by the two definitions of cause, the mental displays exactly the same degree of necessity as the physical. In this chapter, he tackles the three responses he has received to that challenge from Helen Beebe, Peter Kail, and John Wright. As well as providing a rigorous defence of the Old Humean claim that Hume takes the two definitions of cause to exhaust what we can understand about causal necessity, Millican clarifies that what is at issue between the two sides is the intelligibility of thick necessary connections going beyond the two definitions, and demonstrates that there can be, and actually is, substantial variation within the New Humean interpretations.

Sandis, meanwhile, argues that Hume is an *empirical realist* about causation. According to Sandis, what Hume means—and all that he *could*



mean—by ‘necessary connexion’ is the constant conjunction of objects accompanied by the ‘determination of mind’ that leads us to infer the latter from the former. This interpretation is therefore opposed to both the eliminativist interpretation of Hume as someone who *denies* the existence of necessary connexions, and the ‘New Hume’ who is a *sceptical realist* about necessary connexions (such connexions may well exist, but we are never directly aware of them). Like Millican, Sandis takes Hume’s two definitions to be semantic in intent, and central to his argument is an account of Hume’s view of meaning, according to which the meaning of a term consists in its ‘signification’ or ‘sense’. It therefore rules out the possibility that a term like ‘necessary connexion’ could refer to something, even though we have no intelligible conception of that to which the term refers, as New Humeans insist.

Much of the modern debate about causation derives from Hume. Hennig critically considers an assumption made by Hume, which many subsequent writers on causation share, that causation is a *relation* between distinct existences, be they events, facts, states of affairs, properties, substances, or so on. Although admitting that it is possible to refer to causal processes as relations given some suitable description of the process’s parts, Hennig argues that genuine causation involves causal *processes*.

Agreeing with Hume that relations necessarily hold only between distinct existences, it follows that, if causation is a relation, then cause and effect must be entirely separable; as Hume puts it, ‘the effect is totally different from the cause, and consequently can never be discovered in it’.<sup>4</sup> We can see this claim of Hume’s as a hard-won consequence of mechanism, a consequence which Galileo failed to grasp in his appeal to *horror vacui*, which the Cartesians struggled to reconcile with the dependence of creation on God, and which Hobbes made an important move towards with his rejection of real essences.

Hennig, however, argues that Hume’s argument for the claim that causation is a relation is unconvincing: it does not follow from the fact there is no (discoverable) quality common to all causes that causation must therefore be a relation between distinct existences, as it overlooks the possibility that causation is not a relation but a causal process. Moreover, the stages of a causal process are not mutually independent in the way that the Humean conception of relations requires. Using Hume’s example that water is the cause of sinking, and sinking the cause of suffocation,<sup>5</sup> Hennig argues that it is in fact only reasonable to believe that water is the cause of sinking, and sinking the cause of suffocation, if the boundaries between cause and effect are blurred. For instance, being in water only makes it reasonable to believe that one will sink if one is *already in the process of sinking*. Likewise, sinking only makes it reasonable to believe that one will suffocate if one is *already in the process of suffocating*. Causation only holds between different stages of the same process.



### III. AFTER HUME

Hennig's discussion illustrates Hume enduring influence on our understanding of causation. The chapters in the third part of the volume consider the continuation of the causation debate in the period following Hume.

Breitenbach is concerned with an apparent inconsistency in Kant's views of causation. In the *Critique of Pure Reason*, Kant claims that the principle of causality, according to which '[a]ll alterations take place in conformity with the law of the connection of cause and effect',<sup>6</sup> is an *a priori* law of the understanding, and not something that can be discovered empirically. It is an *a priori* law of the understanding not only in the sense that it makes possible the experience of an objective sequence of successive events, like a ship's moving down a river, but insofar as it is presupposed by the ability to conceive of something as part of an objective time sequence. In the *Critique of Judgement*, however, Kant claims that the principle that nature is governed by mechanical laws—which might seem like little more than an application of the principle of causality—is a merely regulative and subjective maxim. That is to say, the principle that nature is governed by mechanical laws is something that we must assume in trying to understand the natural world, but is not something that we can know for certain.

Are these claims consistent? Breitenbach argues that they are, on the grounds that the principle of causality only requires that we experience the world as subject to causal laws. The principle of causality does not itself determine which set of empirical laws the world conforms to—or even that there is an empirically discoverable set of laws that the empirical world conforms to. That the empirical world is governed by laws, and specifically mechanistic laws that lie within our epistemic grasp, is something that we need to assume as a regulative principle of reflective judgement; it is not something that we can know *a priori*.

The final two chapters examine the reactions to Hume of two philosophers who have been neglected in the recent history of philosophy: Thomas Brown and Lady Mary Shepherd. Both were motivated to write about Hume's theory of causation by the Leslie affair—which occurred just a year after Kant's death—and, in their different ways, respond to the apparently sceptical and atheistic tenor of Hume's writing.

In his contribution to the volume, Psillos argues that Thomas Brown's philosophy of causation represents a subtle synthesis of his predecessors, Hume and Reid. At a metaphysical level, Brown defends a broadly Humean regularity theory of causation. According to Brown, a cause ('as it is in the world') is nothing more than something which regularly and invariably precedes change of a certain kind. As such, Reid's postulation of independent powers—the exercise of which is supposed to produce change, but to which we lack epistemic access—is both confused and unmotivated. It is an unexplanatory hypostatization—like the belief in universals, 'a monstrous



species of realism’—that arises (*inter alia*) from uncritical reliance on metaphorical forms of speech and folk beliefs about causation. Although he agrees with Reid (and, according to at least some readings, Hume) that there are powers, Brown argues for an *identity theory* of powers. According to Brown, a power of an object is nothing more than a particular ‘invariableness of antecedence’;<sup>7</sup> or as Psillos characterises it, to ascribe a power to a particular is to do nothing more than ‘to place this substance in a network of regularities . . . that relate this substance with others as cause and effect’ (p. 227)

Brown nevertheless diverges from Hume in his account of how causal beliefs arise. Agreeing with Hume that causal beliefs derive from neither experience nor reason, Brown argues that causal beliefs are what Reid classified as ‘intuitive’. For Brown, Hume’s account of the production of causal beliefs is neither psychologically plausible nor epistemologically effective. Instead of explaining causal beliefs in terms of Hume’s impression of reflection, Brown argues that causal beliefs are simply the immediate and irresistible products of intuition, and as such the self-evident Archimedean starting points of rational inquiry.

This combination of Humean metaphysics and Reid’s commonsense philosophy allows Brown to combine a regularity theory of causation with the existence of singular causal beliefs: beliefs that do not depend on experience of constant conjunction.

Bolton considers a different response to Hume, that of Lady Mary Shepherd. Like Brown, Shepherd’s interest in Hume was ignited by the Leslie affair; unlike Brown, Shepherd defended a thoroughgoing anti-Humean theory of causality. Central to Shepherd’s account, the overarching aim of which is a defence of deism, is what Bolton calls the ‘Causal Principle’: that everything that begins to exist has a cause. Shepherd’s argument for this principle, like Hume’s argument against it, is embedded within a particular theory of mind; and as Bolton points out, the dispute between Shepherd and Hume over causality can effectively be seen as a way of determining the explanatory success of their respective theories of mind.

Rejecting a purely sensational theory of perceptual experience, Shepherd argues that we perceive events as being such that there is causal connection between what begins to exist and an enduring object (or objects), such that what begins to exist inheres in the enduring object. To the extent that we can conceive of something beginning to exist that has no cause, this is only because we do not conceive of the concomitant circumstances necessary for the event. In claiming to be able to imagine the sun’s failing to rise tomorrow, for instance, Hume simply ignores many of the circumstances that would be necessary for this event to occur.

Underlying this argument for the Causal Principle is a conception of causality that is importantly different from the conception attacked by Hume. For Shepherd, cause and effect are simultaneous, not successive, and effects are contained within their causes as are the elements within a mixture.



This defence of the Causal Principle has a number of striking anti-Humean ramifications. It supports a rational account of causal inference; it entails the impossibility of change in the course of nature (though, through a separation of metaphysics and epistemology, it is consistent with ignorance of future courses of events); and the theory of induction is generalised to mathematics and geometry to yield a unified account of *a posteriori* knowledge of necessary truths.

#### IV. CONCLUDING REMARKS

Any period of the history of philosophy as complex as the modern period can be understood by following any one of a variety of threads through the rich tapestry. The chapters in this volume, when considered as a whole, can help us see how the debate about causation arises not only from the success of the new mechanistic science but also a deep philosophical tension. While effects will depend in some way upon their causes, the causes, it seems, must be independent of their effects; philosophers consequently sought the true causes among the true substances, those paradigms of independent existence. But this independence is hard to reconcile with the necessitation of the effect by the cause: the necessary connexion between cause and effect appears to make for a mutual relation which cannot be explained, in a mechanistic world picture, by appeal to the teleology of Aristotelian final causes.

So in the pre-Humean, Cartesian phase we see a struggle with the tension between a substance metaphysics and a mechanistic necessity, with Hobbes beginning to lead the way to the rejection of the former in favour of the latter. In Hume we find an explicit rejection of substance metaphysics but also an insistence on the distinctness of cause and effect, which leads him to reject any notion of necessary connection thicker than mere regularity, or if the New Hume interpreters are right, to at least push it beyond the proper realms of science.

While we may wonder whether Hume's alternative to substance metaphysics really removes the tensions in seventeenth-century thinking about causation, what we see in the post-Humean phase is an awareness that Hume has given an inadequate explanation of why we continue to seek to make the world intelligible in terms of causal necessities. Hume's explanation in terms of his theory of mind is weak precisely because of the weakness of that theory of mind, and in the post-Humean discussion we find various alternative theories of mind put to the task of explaining our causal reasoning.

While the chapters in this volume are primarily aimed at improving our historical understanding, it should also be clear to the philosophically informed reader that most of the issues which these historical disputes turned upon remain unresolved and thus that an exploration of these debates will help further our own philosophical understanding.



## NOTES

1. *Meditations on First Philosophy*, in *Œuvres de Descartes*, edited by C. Adam and P. Tannery, 11 vols. (Paris: Vrin, 1964–74), VII, 49 (abbreviated as AT), and *The Philosophical Writings of Descartes*, translated by John Cottingham, Robert Stoothoof, and Dugald Murdoch, 2 vols. (Cambridge: Cambridge University Press, 1985), II, 33 (abbreviated as CSM).
2. Robert C. Sleigh, Jr., “Leibniz on Malebranche on Causality”, in *Central Themes in Early Modern Philosophy: Essays Presented to Jonathan Bennett* (Cambridge: Cambridge University Press, 1990), 161–93.
3. AT XI, 39, CSM I, 93.
4. David Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, edited by L. A. Selby-Bigge and P. H. Nidditch (Oxford: Clarendon Press, 1974), IV, I, 29.
5. David Hume, *A Treatise of Human Nature*, edited by L. A. Selby-Bigge and P. H. Nidditch (Oxford: Clarendon Press, 1978), I.iii.8, 104.
6. *Critique of Pure Reason*, translated by Norman Kemp Smith (London: Palgrave, 1929, first published 1781/1787), B232.
7. Thomas Brown, *Inquiry into the Relation of Cause and Effect* (Andover: Flaggand Gould Printers, 1822), 21.



# 1 Galileo

## Reflections on Failure

*David Wootton*

### I. GALILEO'S THEORY OF THE TIDES

Galileo was one of the greatest scientists of all time.<sup>1</sup> As a result of his determination to publish his results, he faced trial by the Inquisition and condemnation. It is scarcely surprising that most accounts of Galileo are fundamentally uncritical. An older literature—from Duhem to Koestler—which argued that he had achieved much less than he claimed now receives little support.<sup>2</sup> A newer literature—from Feyerabend to Biagioli—insists that his work consisted of a series of public performances, that it was geared to obtaining patronage and status; but this literature insists that this is true of science in general, and so blunts its critique of Galileo—a critique which is, in any case, largely misconceived.<sup>3</sup> My argument here is that understanding Galileo requires us to think about his failures, and that those failures are evidence of his confused thinking on the subject of causation. My conclusion is a cautionary one: consistency alone would not have rescued Galileo. But it is clear from the responses of his peers that Galileo's mistakes were unforced. They resulted from a commitment to Copernicanism that went beyond reason—as interesting commitments often do.

Very early in his career, in 1592, Galileo became a Copernican.<sup>4</sup> At the time (roughly fifty years after the publication of *De Revolutionibus*), a public commitment to Copernicanism was extremely rare. Only one competent mathematician—Thomas Digges—had published in defence of the Copernican system. Tables which employed Copernican mathematical techniques were not uncommon—but the astronomers who produced them, such as Giovanni Magini, were usually firmly opposed to heliocentrism. Galileo's conversion, which was to have a determining influence on his intellectual activities throughout the rest of his life, was grounded in a basic discovery. In collaboration with his patron, Guidobaldo del Monte, he had carried out an experiment which established the parabolic path of projectiles.<sup>5</sup> Armed with this new knowledge, Galileo was convinced that he could rebut the standard arguments against Copernicanism, which were *physical* not astronomical: why, if the earth is moving under our feet, do we not fall over? Why, if an object is dropped from a high tower, does it not land to



the west of the tower's base? Why does a cannon shot not travel further if fired to the west than if fired to the east? Galileo's project for a new physics, which culminated in the *Two New Sciences* of 1638, was designed to destroy the physical arguments against Copernicanism; his new astronomy—the discovery of the mountains on the moon, of the moons of Jupiter, of sunspots, and of the phases of Venus in a few brief months between the winter of 1609/10 and the winter of 1610/11—decisively refuted Ptolemaic astronomy, but left the alternative system of Tycho Brahe relatively undamaged. What Galileo lacked, in order to prove Tycho wrong, was proof that the earth moved. Attempts to prove this by, for example, measuring stellar parallax had proved uniformly unsuccessful.

This brief summary enables us to place Galileo's two great failures in context. The first is well known. From an early date (our earliest evidence dates to 1595), Galileo decided that the only possible explanation of the tides must be that they were a consequence of the motion of the earth.<sup>6</sup> According to Copernicus, the earth both orbits the sun in the course of a year and rotates on its axis in the course of a day. The combination of these two movements means that any point on the surface of the globe (apart from the North and South poles) speeds up and slows down in the course of every twenty-four hours—for twelve hours its speed is the result of rotation being added to orbital motion; while for twelve hours its speed is the result of rotation being subtracted from orbital motion.

Galileo's theory of the tides faced substantial difficulties. Galileo developed and adapted it to try to explain why there are two tides a day, not one; and why the heights of tides vary on an annual and on a lunar cycle. He published his final version of the theory in his *Dialogue Concerning the Two Chief World Systems* of 1632.<sup>7</sup> His original intention—thwarted by the Roman censors—was that his book should be entitled *Dialogue on the Tides*, and his fundamental argument in that book was that, since only the double movement of the earth can account for the tides, the Copernican account of the universe must be correct. Since any explicit defence of Copernicanism had been banned by the Catholic Church in 1616, Galileo's argument laid him open to the charges that rapidly followed publication and that inexorably led to his trial and condemnation.

The crux of the matter is this: Galileo dismisses all alternative accounts of the tides (accounts, for example, which explained them in terms of some sort of attraction exercised by the moon) as fantastical—*only a mechanical explanation will do*. Indeed, he sought to construct a machine which would replicate the earth's movement and cause tidal movement in a miniaturised ocean. Unfortunately, Galileo was wrong—no mechanical system can adequately account for the movement of the tides as they occur on this planet.<sup>8</sup> Despite the great sophistication of Galileo's theory, he had no explanation for the fact that the time of high and low tide changes in a regular fashion from day to day—as was immediately pointed out by his friend and rival Giovanni Battista Baliani.<sup>9</sup>



There is no evidence that Galileo's tidal theory ever convinced *anybody*. It is worth pausing to consider what would have happened if he had not committed himself to it and had not decided to pursue it in the face of every obstacle. Had he chosen to go on debating Copernicanism after the condemnation of 1616, but had not introduced his theory of the tides, he would have been able to argue that the Ptolemaic system had been destroyed by the discovery of the phases of Venus.<sup>10</sup> And he would have been able to argue that the Copernican system was in certain respects more plausible than the Tychonic system—the earth, for example, shines by reflected light, and all the other bodies that shine by reflected light are in movement; the sun shines by its own light, and all the other bodies that shine by their own light appear to be fixed.<sup>11</sup> But he would not have been able to claim that he could prove the truth of Copernicanism. There would have been no trial in 1633, no condemnation of Galileo, and no formal judgement that Copernicanism was heretical. As a result, there would have been far greater intellectual freedom in Catholic countries over the course of the subsequent two centuries (Copernicanism was not officially permitted by the Catholic Church until 1835). The scientific revolution, which flourished primarily in northern Europe, out of reach of the Inquisition, might well have advanced more rapidly and more generally. There might have been an Italian or even a Spanish Newton.

Galileo's commitment to his theory of the tides was a major intellectual error. It appears to have resulted first from an overenthusiastic determination to find arguments in favour of Copernicanism, and second from a dogmatic insistence that in physics causal explanations have to be mechanical explanations and cannot involve action over a distance. This insistence flew in the face of the acknowledged difficulty of providing, for example, a mechanical account of magnetism—a subject Galileo had studied with great care over many years.<sup>12</sup>

## II. GALILEO AND THE HORROR VACUI

The story I have told so far appears straightforward. Galileo was committed to a mechanistic physics, and this commitment led him astray. It was, indeed, responsible for his greatest failure. But I want now to argue that Galileo's *lack of commitment* to mechanistic explanations becomes apparent as soon as we look at Galileo's other great failure, his failure to resolve basic questions about the production of a vacuum. I hope to convince you of a basic paradox: Galileo was *too* committed to mechanism, and at the very same time, he was *insufficiently* committed to mechanism. Resolving this paradox will not be easy.

Even before Galileo became a Copernican, he was opposed to the Aristotelian view that some substances are *heavy* and descend towards the centre of the earth, and others are *light* and rise towards the heavens.<sup>13</sup> Galileo, whose early science was modelled on Archimedes, insisted that heaviness



and lightness were relative and not absolute terms—light objects float upwards not because they are light but because they are less dense than heavier objects which crowd them out. Relative density (or specific gravity) thus explains why bubbles of air rise in water, or why wooden sticks descend when released in air. In 1611 Galileo defended this Archimedean account against his Aristotelian opponents, who argued that it is shape not relative density that determines whether objects float—they held that ice is denser than water but floats because of its flat shape; while Galileo argued that the fact that ice floats is proof that it is less dense than water.<sup>14</sup>

Galileo was thus committed to the view that air has weight; that we live in an ocean of air; and that it presses on us from all sides. He carried out experiments to determine the weight of air by measuring containers with more or less air in them.<sup>15</sup> Galileo's study of flotation had led him to explain how, within a confined space, a small weight of displaced water could support the far greater weight of a floating body—a version of the principle of the lever, which provides the theoretical foundation of the hydraulic press. In the Arsenal at Venice he had seen ships which were robust when floating in the water, but which had to be supported by shoring when on dry land—when floating they were held together by the pressure of the water, while on dry land they would have broken under their own weight.<sup>16</sup> Galileo was thus an expert on how pressure works.

Moreover, even before Galileo became a Copernican, he had been prepared to accept the possibility of a vacuum—in contrast to the Aristotelians, who argued that there could be no such thing. By 1630, if not earlier, Galileo reasoned that a pump cannot raise water more than approximately thirty-four feet because if the column of water is higher than that it breaks and a vacuum appears. He thus held that there was nothing mechanically difficult about producing a vacuum; all one needed was a well where the water level was more than thirty feet below the surface, and a pump.<sup>17</sup> Moreover, his friend Sagredo had pointed out to him that there was a vacuum in the top of every thermometer—Sagredo's thermometers were made out of wide tubes, so that the vacuum space was substantial enough for him to conduct simple experiments, and to conclude that sound did not cross the vacuum.<sup>18</sup>

Yet Galileo—who had conducted extensive experiments with pendulums, floating bodies, magnets, and projectiles—never conducted any vacuum experiments. He did not explain the limit on how high a pump could raise water in terms of air pressure; instead, he related it to theories regarding the breaking strength of materials. Take a rope: if you lift it higher and higher into the air, eventually it will break under its own weight. This, Galileo argued, is what happens in a high-rise pump. Normally some force holds the water together; but when the weight of the water passes a certain limit, that force is overcome. What force is this? It is obvious that anything can cut through water, and that the fluidity of water implies that its parts slip past each other. So water cannot be held together as a rope is. Rather, it is held together in the same way that two polished slabs of marble



are held together if one is placed on top of the other—lift the top one and the bottom one comes up with it. The force holding the column of water together and holding the two marble slabs together is the resistance to the production of a vacuum—Galileo maintained that it was easy to produce a vacuum, but to do so one had to overcome a measurable *horror vacui*, a mysterious force that acted to prevent vacuums from coming into existence. Since this force was rather powerful, much of the strength of materials, Galileo reasoned, derived from it.<sup>19</sup>

It is important to see what has happened here. Galileo had to hand all the intellectual tools required to account for the failure of suction pumps in terms of air pressure. This would have been a mechanical Archimedean explanation. Indeed, such an explanation was carefully laid out for him by Baliani in a letter of 1630—there, Baliani seems to be tactfully suggesting that such an explanation would work better than the explanation Galileo is proposing, although he acknowledges that he would expect the force of air pressure to be greater than it appears to be (this for the simple reason that Boyle's law had yet to be discovered—it would not have occurred to Baliani that the weight of air in a given volume diminishes as one rises through the atmosphere).<sup>20</sup> But instead, Galileo adopted a nonmechanical explanation in terms of an occult force, the *horror vacui*. This explanation required an appeal not to mechanically efficient causes, such as weight, but to an Aristotelian final cause—when it comes to the vacuum (and only when it comes to the vacuum), nature appears to be purposive. It is true that this explanation seems to have been a placeholder for a more sophisticated explanation—Galileo went on to argue that all material objects contain infinitesimal vacua, so we do not need to assume that the cause of the *horror vacui* is a vacuum that does not yet exist.<sup>21</sup> But he never explained how these infinitesimal vacua might act to prevent the creation of a substantial vacuum.

Within a year or so of Galileo's death in 1642, Roman philosophers set out to prove that he had been wrong to claim that it was easy to make a vacuum. They created a tube sealed at one end which was more than thirty-four feet long, filled it with water, and upended it, open end down, in a butt of water; to their astonishment it turned out that there appeared to be a void at the top of the tube. Galileo had been vindicated. Galileo's disciple Torricelli suggested conducting the same experiment with mercury rather than air, thus creating an apparatus which was much easier to manipulate. In the resulting Torricellian space at the top of the tube it was possible to include another mercury-filled tube, and to show that no column of mercury could be supported within a void. Thus it was apparent that the presence of air was the precondition for supporting a column of water or mercury within a container sealed at one end—which convincingly showed that air pressure was the cause of the phenomenon. Pascal went on to show with the Puy de Dôme experiment (in 1648) that there was further good evidence to support this hypothesis. Thus Torricelli's tube was transformed into a barometer.<sup>22</sup>



*Torricelli and Pascal had no conceptual tools at their disposal that were not available to Galileo—indeed, key elements of a correct account had been laid out for Galileo by Baliani.* It should not be surprising that Pascal's account of his vacuum experiments often seems to echo Galileo's account of floating bodies, so closely is he dependent on principles familiar to Galileo—indeed he may be directly following Galileo's text. After all, as Baliani had stressed, the crucial step is to recognise that air pressure functions in exactly the same way as water pressure. Had Galileo conducted experiments on vacuums he would soon have had to abandon his *horror vacui* thesis (even if he had examined Sagredo's thermometers he would have been bound to acknowledge that there was no evidence of a *horror vacui* at work), and he would have had, in Baliani's account of air pressure, an alternative ready to hand.

When Copernicanism was condemned in 1616, Galileo could easily have turned to experimentation on vacuums, whose scientific significance was indisputable—for someone engaged in the enterprise of refuting Aristotelian philosophy, as Galileo was, the vacuum was clearly of strategic importance. Instead of publishing the *Dialogue* he could have carried out the experiments later performed by Torricelli, Pascal, and even Boyle. Overconfident that he could prove the truth of Copernicanism by relying on a mechanical account of the tides, he continued to concentrate on cosmology; somewhat carelessly, he then allowed himself to rely on a nonmechanical account of the *horror vacui*.

How can we explain the paradox that Galileo was both too much a mechanist and too little a mechanist? The only answer I can see is that Galileo was a mechanist when mechanism served the cause of Copernicanism. Mechanism was simply a tool to achieve a desired effect; had Galileo's commitment to it gone deeper he would never have been satisfied with his account of the mysterious force required to produce a vacuum.

As for Galileo's commitment to Copernicanism—Galileo had plenty of reasons for preferring Copernicanism to the available alternatives; but he overstated his arguments in support of it because his commitment to Copernicanism went beyond reason. In the end, Galileo preferred the idea that the universe had not been made for man over the idea that the universe had been made to provide mankind with a home; this fundamental psychological commitment underpinned his work on cosmology, while there was no comparable internal force driving forward his work on vacuums.<sup>23</sup>

### III. GALILEO AND THE ESCAPE FROM ARISTOTELIAN CAUSATION

Let us turn from these two specific examples, Galileo on tides and Galileo on the vacuum, to ask more generally: what was Galileo's view of causation? The question seems simple but is actually difficult. Galileo has very



little to say about causation—but he never tells us why he thinks causation is unimportant. Our question is worth asking, however, because it proves an excellent way of approaching an even more basic question: what sort of science was Galileo's? To deal with this pair of questions we need to overcome a series of obstacles.

The first problem is this: from the beginning of his career until its end, Galileo was opposed to Aristotelian accounts of how nature works. On the other hand, he insisted, late in life, that he had never been opposed to Aristotle's logic.<sup>24</sup> Would he have included Aristotle's theory of causation within Aristotle's logic? There is nothing to suggest that he would, and indeed much of Galileo's science assumes that Aristotle's account of how nature works is inseparable from his theory of causation. Galileo's new science implies a new approach to causation, although this new approach is generally implicit rather than explicit.<sup>25</sup>

The second problem is this: historians and philosophers of science tend to believe that a preoccupation with causation in nature, characteristic of classical and medieval science, has been replaced in modern times by a search for laws of nature. In this story the seventeenth century is a tale of two philosophies: on the one hand there is the mechanistic philosophy, which brings with it a backward-looking preoccupation with causation; on the other, there is the new experimental method, which is accompanied, in the programmatic statements of the Accademia del Cimento and the Royal Society, by a deep scepticism about the Aristotelian and Cartesian rush to identify causes. In any case, along comes Newton, and causation ceases to be important. Here Galileo appears to be an exception: although he did most of his decisive work in the first dozen years of the seventeenth century, although in the *Assayer* he laid the foundations of later mechanistic philosophy, although he presents himself as being first and foremost an experimental scientist, he seems uninterested in causation. Moreover, Galileo is heralded as the discoverer of two fundamental laws of modern physics, the 'law' of the pendulum (I place 'law' in quotation marks, because the law of the pendulum, as presented by Galileo, is inaccurate) and the law of the acceleration of falling bodies. Looked at in this light he seems to be a precursor to Newton and a founder of 'modern' science—Stillman Drake was happy to present him in this light. It is therefore puzzling indeed that he never writes about laws of nature, although the concept was surely one with which he was familiar.<sup>26</sup> Galileo will never find his place in a story which assumes that first there were causes (whether Aristotelian or mechanical) and then there were laws. Looking at him helps us see that there is something wrong with this story.

The third problem is this: Galileo worked in a number of different fields, although the bulk of his work bore on issues related to the truth or falsity of Copernicanism. Galileo's handling of questions of causation varies, as one would expect, from field to field, and so it would be a mistake to look for one view of causation in his work, or to impose an artificial unity upon his



thinking. There is no more a single Galilean view of causation than there is a unitary Galilean science.

These difficulties mean we have to tread carefully, but they do not represent an insuperable obstacle. And if we persist we will end up with a much better understanding of Galileo's science, and a better sense of some of the ambiguities attached to the idea of causation. More important still, we will acquire a new understanding of the origins of modern science.

Let us begin with Aristotle. According to Aristotle all natural processes are to be understood in terms of causation, and a causal process requires the copresence of four types of cause: formal, final, efficient, and material. Galileo avoids, presumably deliberately, using these categories. It seems clear that he does not believe (with the solitary, awkward exception of the *horror vacui*) in final causes in physics and astronomy—if God has purposes they are concealed from us. It also seems clear that he does not believe in forms, only in arrangements of matter—it seems very likely that he is a materialist. We could reconstruct many of his arguments in terms of efficient and material causes, but that is only to say that we could reconstruct them in terms of a mechanistic theory of causation. If we set out to do this, though, we would be misunderstanding what Galileo is trying to do—as I hope to show in what follows.

Here is a simple thought experiment: Put two weights in the scales of a balance. The heavier one will go down; if neither goes down, then the two weights are equal. We can describe this as a causal process or as a mathematical relationship. We can go on to the yardarm or the lever, and discover the ratios of weight, speed, and distance moved that are involved in the balancing and moving of unequal weights and forces. We can treat complex problems where different weights are distributed at varying distances along the arms of a balance and attached in different ways. Galileo's intellectual career begins and ends with problems of this sort.

Two things are important about the way in which Galileo handles them. First, he deals in pure mathematical abstractions—he is not interested in real balances, or actual levers, as these are necessarily imperfect. They stick, they flex, they wobble. Second, he aspires to mathematical proofs, but he is often happy with thought experiments. Thus he will provide an alternative drawing of a problem in order to show why a claim he makes must be right; and will invite the reader to imagine manipulating a balance by hanging weights from it in various ways (hanging weights from both ends and from the middle, dividing them and hanging them back in place) in order to convey a mathematical relationship between distance from the fulcrum and force exerted.<sup>27</sup> Galileo's model in all this is Archimedes, and we could say that these Archimedean relationships are both mathematical and thought-experimental (the thought experiments drawing on our undertheorised experience of the world).<sup>28</sup> The same ways of thinking go into his work on mechanics. In this world there are causes (heavy weights overbalance light weights in the scale, and there is no doubt that there is a



cause for the ascent of the light weight) but the only causal mechanism is the principle of the lever (which provides the key to an understanding of all the simple machines).

How are we to understand the type of knowledge involved in Galileo's mechanics? Thought experiments involve the manipulation of conceptual models—they appeal to a manipulative theory of causation even if they do not involve actual experiments, but ask us simply to reanalyse our previous experience. The principle of the balance is a stipulated premise, comparable to the claim in geometry that a line is the shortest distance between two points. With this premise it becomes possible to formulate a deductive science of mechanics. How do we know that two weights in the pans of a balance are equal if the balance is in equilibrium? The argument is entirely circular: the weights are equal because the balance is in equilibrium; the balance is in equilibrium because the weights are equal. We define equal weights as weights that balance. If we want to go beyond that we have to appeal to direct experience—they feel equal if we hold them in our hands.

Galileo hoped to develop other sciences on this model. When he first tackled the problem of motion, in the early 1590s, he refuted the Aristotelian claim that heavy weights fall faster than light weights by a simple thought experiment—tie two one-pound weights together with a rope. Will they now fall twice as fast as one one-pound weight? If so, which of the one-pound weights will pull the other one down? If the rope remains slack, then clearly they are falling as if they are two separate one-pound weights.<sup>29</sup> Galileo aimed to replace the Aristotelian theory by the claim that rate of fall was determined by relative density or specific gravity. He conducted some experiments which failed to illustrate this theory—but this did not unduly worry him, as he hoped to show that the relationship between speed of fall and specific gravity was a *necessary* relationship.

Specific gravity, as we have seen, also provides the key to Galileo's account of flotation. According to Aristotle, some objects are naturally heavy, and some naturally light—they ascend while heavy objects descend. Flotation, on the other hand, is determined not by weight but shape—broad objects float, narrow objects sink. Ice is heavier than water, but floats because of its shape. Galileo had no time for this theory. In his view objects sink if their specific gravity is greater than the specific gravity of the medium they are floating in. They are extruded or crowded out by the descent of the denser material, and this is generally true—air is not light and keen to ascend, but heavy and seeking to descend, but it is crowded out by denser objects. Ice floats because it is lighter, not heavier than water. Specific gravity thus provides a causal explanation for speed of fall (until Galileo abandoned this theory in 1592) and for flotation. Is relative density a formal, efficient, or material cause of flotation? The answer would have to be that it is all three, or rather that the question is meaningless, or trivial. What matters is that the question of flotation has been reduced to a mathematical ratio, just as the work done by a lever is defined by the ratio between the distance from



the fulcrum to the force and the distance from the fulcrum to the load. In an Archimedean science, causation is like the windscreen of your car—unless something goes wrong, you look through it not at it.

Galileo taught a number of subjects apart from mechanics—astronomy/astrology, optics, fortification. All of these work on geometrical principles. Thus in Ptolemaic and Copernican astronomy, the heavenly bodies move in circles. In optics, light travels in straight lines and angles of incidence and angles of reflection are equal. In warfare, guns shoot straight, so that the task of calculating fields of fire, and of working out how to bring flanking and raking fire to bear on attackers, is a simple geometrical problem. Because he does not have access to algebra or calculus, Galileo relies on geometry to address problems such as those of the acceleration and average speed of falling bodies. When he says that the book on nature is written in lines, circles, squares, and other geometrical figures, he means that geometry provides the tools for building his new science.

This is true to a much greater extent than is generally recognised. The first of Galileo's great discoveries in astronomy, made in December 1609, is the presence of mountains on the moon. You cannot *see* the mountains on the moon. What you can see are the shadows they cast, and the fact that as the sun rises over the lunar landscape it catches the peaks of the mountains before it shines into the valleys. The phenomena Galileo is discussing were regularly taught within perspective, a subject taught to artists by mathematicians, and one with which Galileo will have been familiar, even if he never taught it himself. Artists had learnt how to convey three dimensions on a two-dimensional surface; Galileo, looking at a two-dimensional image of the moon, is able to read it as if it had been painted by an artist, and recognise it as an image of a three-dimensional object.<sup>30</sup> Because the principles of perspective were generally understood, and had nothing to do with telescopes or Copernicanism, Galileo's claim to have discovered mountains on the moon quickly ceased to be contentious and became the new orthodoxy. Again, we could give a causal explanation—shadows are caused by the fact that light travels in straight lines. But since we all already know this, the causal explanation is not at all interesting—we look through it to the novel claim that there are mountains on the moon.

We find exactly the same phenomenon with Galileo's other major astronomical discoveries—they all depend on noncontentious causal claims derived from elementary principles of perspective. Thus Galileo discovers that Venus has a full set of phases, and argues that this demonstrates that Venus orbits the sun not the earth—the claim is irrefutable if one accepts that Venus is illuminated by the sun, and that half of its sphere is illuminated at any one time. (On the Ptolemaic system no more than half of Venus as seen from the earth could ever be illuminated.) By a more complicated argument, Galileo demonstrates that sunspots are on or very near the surface of the sun—he shows that the way in which they are foreshortened as they approach the edge of the sun as the sun rotates requires that they be



surface phenomena.<sup>31</sup> Again, the mathematical principles underlying foreshortening are too well-known to be interesting in and of themselves. And he argues that the new stars he has found at varying distances from Jupiter must be moons orbiting the planet because he can take for granted a model according to which heavenly bodies are engaged in various types of (at least roughly) circular motion—he therefore does not even have to discuss the possibility that they are shuttling back and forth between the surface of the planet and deep space because there is no analogue for such a possibility, while our own moon provides an analogue for the behaviour of the moons of Jupiter.

I argued earlier that it may be true that the heavier weight causes an arm of the balance to descend, but that is in the end a circular argument because we define a heavier weight as the one which descends. Similarly, one might say that it follows automatically from the claim that light travels in straight lines that there are shadows (what is puzzling is not the presence of shadows but the fact that they are not pitch-black—here we have to consider the way in which the sun's light is refracted in the atmosphere), and it follows automatically from the claim that Venus has a full set of phases that it orbits the sun. Galileo is not interested in describing these systems as causal because the causal mechanism is too obvious to be interesting.

Much of Galileo's science is thus causal—but the causes are trivial and transparent. What matters is that he is able to save or solve phenomena by reducing them to mathematical principles. This is the essential background for Galileo's work on a new theory of fall. This began in 1592, when, as we have noted, he discovered through experiment that the path of a projectile is symmetrical, and seems to have been largely completed soon after 1604. Galileo's ambition was to produce a deductive theory of the acceleration of falling bodies, although he had difficulty demonstrating his initial premise—he was still revising his argument after the publication of *Two New Sciences* in 1638. Galileo carried out sophisticated experiments to confirm his theory, but he was quite clear that his approach differed from that of Baliani, who produced an account of fall more or less identical to Galileo's on the basis of empirical evidence. Galileo has no explanation of *why* falling bodies accelerate; he does, on the other hand, have a proof to show that if you mount a circle vertically, a body will take the same time to fall along any chord which touches the lowest point of the circle. In other words, he can establish a series of mathematical relationships relating to acceleration, as he can relating to machines or the foreshortening of sunspots. In this case there is no 'cause' to look through—but this hardly seems to concern Galileo, because the phenomenon he is looking at turns out to be mathematically legible (the path of a projectile, for example, turns out to be a parabola). Galileo slips from a causal to a noncausal science almost without noticing it.

Where Galileo's deductive science broke down he was forced to turn to empirical confirmation: thus he hoped to show that the time taken by any



pendulum of a fixed length to swing must be a constant, no matter what the arc of swing, but he could establish this empirically (or so he believed) but not deductively—the problem was eventually solved by Huygens. In Galileo's hierarchy of values, deductive sciences like geometry come first; second come observational sciences grounded in universally acknowledged mathematical principles (such as that protuberances cast shadows); and third come inductive generalisations and experimental 'proofs'. It is this hierarchy which matters to him, not the issue of causation. Despite his own use of experiments and his admiration for William Gilbert's *De Magnete*, a purely experimental science is, in Galileo's eyes, always third best.

#### IV. MECHANISM AND INTERVENTION

Now let us consider two complications. When I look through the windscreen of my car I do not only look through the glass; I also look past the squashed bugs which obscure my view. Galileo may generally adopt a windscreen approach to causation, but on the rare occasion when his eyes happen to focus on the windscreen, there are two things that he sees. In the first place, Galileo in the *Assayer* presents the first account in modern natural philosophy of a mechanistic system (as we would call it) founded on a distinction between primary and secondary qualities.<sup>32</sup> Galileo's suggestion is that extension is primary, but colour, sound, smell, and taste are all caused by waves or particles which are perceived as colours, sounds, smells, and tastes—these appearances are in the observer not in the object, as the sensation of being tickled is in the person being tickled, not in the feather with which he is tickled. Shape, on the other hand, really is in the object—this is a conceptual, and not just an observational, truth. Galileo thus proposes that smells are the result of particles entering the nose, and different smells are caused by particles of different sizes and shapes. Particles are real; smells are subjective.

It is a little difficult to know whether all later discussions of primary and secondary qualities derive, directly or indirectly, from Galileo, or whether they derive, as Galileo's did, from a direct encounter with the classical texts dealing with atomism. What is clear is that Galileo had sketched out a mechanical philosophy, and that later philosophers and scientists, such as Descartes and Boyle, were working within parameters that he had recognised and perhaps established. The mechanical philosophy requires a strong theory of causation, in that, for example, smells have to be caused by something that has apparently no connection with smell, and certainly cannot be said to follow automatically from a definition of what a smell is. Galileo is thus genuinely interested in the idea of a science which will discover new, hidden causes (I deliberately use the word 'hidden' rather than 'occult'—Gilbert has no account of what causes magnetism, with the result that the cause remains occult, rather than being a hidden cause that



can be uncovered). He evidently thinks there are reasons for thinking that the mechanical philosophy is the only coherent account of how the universe works. But at no point does he claim to have made advances in this new science. Galileo may believe in the mechanical philosophy, but he does not practice it. And (perhaps surprisingly) he never describes the universe as a machine or compares it to a clock, so he never identifies himself as a philosopher of mechanism.

So when Galileo philosophises about causation he distinguishes primary and secondary qualities: if he looks at the windscreen rather than through it, this is what he sees. But this is not the only conception of causation that Galileo has. In 1611, as we have seen, Galileo became engaged in a dispute about the cause of flotation. His opponent, Lodovico delle Colombe, claimed to have refuted Galileo's relative density theory of flotation by floating an ebony chip on water. Galileo was thus obliged to engage in a series of experiments. He showed that it was not—as Colombe claimed—the chip's shape that determined whether it floated or not. Galileo discovered that he could float a needle on the water's surface if he placed it there very gently—he was exploring what we would now call the phenomenon of surface tension. And Galileo showed that he could determine whether an object floated or sank simply by altering its relative density. He took a ball of wax and added slivers of metal to it one by one—as a result he found he could make the ball rise rapidly from the bottom, or sink rapidly from the top, or rise or sink slowly, or float, neither sinking nor rising, wherever it was placed (the Galileo thermometer, which appears to have been invented by Galileo's disciples, exploits this principle).

In his notes as he carried out these experiments, and in the book he wrote describing them, Galileo defines a cause as something which has a relationship to an effect such that if you take away the cause you take away the effect: '*Causa è quella, la qual posta, segue l'effetto; e rimossa, si rimuove l'effetto.*'<sup>33</sup> In an excellent recent article Steffen Ducheyne has described this as Galileo's interventionist theory of causation, and has argued that this was a novel mode of thinking which represented an important contribution to the development of a scientific method.<sup>34</sup> There is only one flaw in Ducheyne's argument—there is nothing novel in Galileo's interventionist theory of causation. The very same view of causation is presented in a work well known to Galileo, Alhacen's eleventh-century work on optics, and it seems evident that Galileo was simply quoting Alhacen.<sup>35</sup> Insofar as the interventionist notion of causation is a natural offspring of the experimental method, and Alhacen had a good grasp of the experimental method, this is exactly what one would expect. Moreover, the familiarity of this definition of cause helps explain why Galileo makes so little fuss about it. Galileo was interested in stressing the importance of his own original contributions, not in banging the drum for others, particularly when their work was already well-known.

It is worth stressing, however, an important difference between Alhacen's and Galileo's interventionist notion of causation and Aristotle's. For



Aristotle, a father is the (formal) cause of his son; but once procreation has taken place the father can die, and the son will continue to live—indeed, he may even be born after his father's death. Similarly, a saw is the (efficient) cause of a table; but once the table is made the saw can be given away without unmaking the table. Alhacen and Galileo are interested in causes that are copresent with their effects: if one alters the relative density of an object one immediately alters its relationship to the water in which it sinks, swims, or floats.

Galileo's interventionist notion of cause appears to have had one important consequence that has not been noticed. Right up until his experiments on floating bodies, Galileo was practicing and defending astrology; immediately thereafter he seems to have lost faith in what was still a highly respectable science. And in the *Dialogue* he suggests that the problem with astrology is that you cannot test it by revising the parameters and seeing how that alters outcomes. Astrology claimed a causal relationship between certain conjunctions in the heavens and certain types of events on earth; an interventionist theory of causation required that one be able to test this relationship. Galileo seems to have decided that in the absence of an experimental test for the claims of astrology the whole science had to be abandoned as worthless.<sup>36</sup>

## V. CONCLUSION

In this chapter I have argued that Galileo had no coherent and consistent view of causation. If he had one, he would have avoided at least one of his two great mistakes—the mechanical theory of tides and the *horror vacui*. I have acknowledged that Galileo had both a mechanical and an interventionist theory of causation—neither was original to him, but the first, which lies at the root of early modern accounts of primary and secondary qualities, seems to have been more influential than the second. But I have stressed that Galileo always aspired to an Archimedean science within which causation becomes a trivial concept, and attention focuses on mathematical relationships that can be expressed in geometrical terms. Because we are not familiar with this sort of science, the strength of some of Galileo's arguments—for the existence of mountains on the moon, for example—was more apparent to contemporaries than it is to us.

Many years ago, Thomas Kuhn argued that there were two types of science practiced in the early modern period—a deductive, mathematical science, practiced on the Continent, and an inductive, experimental science, which was effectively unique to England.<sup>37</sup> On Kuhn's account, Newton was exceptional in practicing both types of science. My own view is that in this respect Newton was not exceptional but typical: Galileo, Torricelli, Pascal, and Huygens also practiced both types of science, and we could add others (Simon Stevin, for example) to this list. The core enterprise of early



modern science was not *either* experimental or deductive; rather, it was, following the example of Archimedes, both at once.

Kuhn was unable to see this for a simple linguistic reason—he was struck that the French seemed only to take *la physique expérimentale* seriously towards the end of the eighteenth century. But if he had looked at Galileo and Pascal he would have discovered that they use the words for *experience* where we would use *experiment*. Indeed, in French there still is no word that corresponds to *experiment* (or at least, to be precise, there no longer is): any English-French dictionary will offer you *expérience* as the translation for *experiment*.<sup>38</sup> In these circumstances it is hardly surprising that the adjective *expérimental* plays a limited role in philosophy and science. In Italian, *esperimento* appears in the *Vocabolario della Crusca* (1612), but it was an unattractive neologism and Galileo hardly ever used it. Crucially in Italian you can *fare una esperienza* and in French you can *faire une expérience*—experience is thus much closer to experiment in Italian and French than it is in English. In English, by contrast, the logic that anything deliberately studied was an experiment, not an experience, led many philosophers (Bacon and Hume are obvious examples) to write of experiments where we would write of experiences—the subtitle of Hume's *Treatise* is *An Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. Here the meaning of *experimental* is, I think, a deliberate conflation of what the OED calls sense I—relating to experience—and sense II—relating to experiment. Hume makes clear in his Introduction that there can be no deliberate, purposeful experiments in moral philosophy, but only careful observations of human conduct; but he does offer a series of 'experiments' which are combinations of abstract reasoning—thought experiments—and observations of conduct.<sup>39</sup> For Hume, the term 'experiment' bridges what we think of as the gap between experience and experiment, just as 'experience' does for Galileo or Pascal.

Galileo's science was thus experiential in the Italian and French sense or experimental in the Baconian and Humean sense—it is because Kuhn is relying on what is fundamentally a nineteenth-century distinction between experiment and experience that he finds no *experimental* science on the Continent in the early modern period. The driving impulse of the scientific revolution lay not in the discovery of a new type of experimental science but in the conviction that there could be (as in astronomy) a fruitful exchange between mathematical models and practical experience, whether this took the form of observations or experiments. Both produced what were now beginning to be called 'facts'.

Galileo was thus, I think, more typical, and less exceptional, than is generally assumed. But I want to end with a conclusion that seems to me even more striking. When we talk of the scientific revolution the word 'revolution' is a metaphor based on political revolutions. Political revolutions are, I would argue, path-dependent. It is impossible to imagine Jacobinism before moderate constitutional reform has been tried and failed; a



full-scale ideology of equality requires first that obvious inequalities—such as the privileges of the nobility—should be swept away. And there can be no Napoleon until the old order has been destroyed, and the new one has failed in its turn to deliver on its promises.<sup>40</sup> But the scientific revolution was not, I would argue, path-dependent in this way. If Galileo had ceased to defend Copernicanism after 1616 and had turned instead to experimental work on vacuums, the future might have been very different—Galileo might have discovered the barometer and Boyle's law, and Catholics might have remained free to engage in hypothetical debates regarding the merits of Copernicanism. This might have altered the direction and pace of advance in certain key fields, but it would not have radically altered the character of the scientific revolution. Galileo's failings (unlike Robespierre's) were not part of some sort of hidden logic of history. This makes them less significant, and also less excusable.

## NOTES

1. For recent surveys of Galileo's life and work, see John Heilbron, *Galileo* (Oxford: Oxford University Press, 2010) and David Wootton, *Galileo: Watcher of the Skies* (New Haven, Conn: Yale University Press, 2010).
2. Pierre Duhem, *To Save the Phenomena: An Essay on the Idea of Physical Theory from Plato to Galileo* [1908] (Chicago: University of Chicago Press, 1969); Arthur Koestler, *The Sleepwalkers* (London: Hutchinson, 1959).
3. Paul Feyerabend, *Against Method* (London: NLB, 1975); Mario Biagioli, *Galileo Courtier* (Chicago: University of Chicago Press, 1993).
4. The case for this early date is made in Wootton, *Galileo*, 137–56.
5. Jürgen Renn, Peter Damerow and Simone Rieger, "Hunting the White Elephant: When and How Did Galileo Discover the Law of Fall?", *Science in Context* 13 (2000), 29–149.
6. Most recently, R. H. Naylor, "Galileo's Tidal Theory", *Isis* 98 (2007), 1–22.
7. Galileo, *Dialogue Concerning the Two Chief World Systems*, trans. Stillman Drake (Berkeley: University of California Press, 1962), 416–65.
8. Paolo Palmieri has conclusively shown that there was nothing wrong in principle with a mechanical theory of the tides: "Re-Examining Galileo's Theory of Tides", *Archive for History of Exact Sciences*, 53 (1998): 223–375.
9. Galileo Galilei, *Le opere di Galileo Galilei: Edizione Nazionale*, ed. A. Favaro (20 vols in 21, 4th edn, Florence: G. Barbèra, 1968), xiv 343 (abbreviated as OG).
10. On which see Paolo Palmieri, "Galileo and the Discovery of the Phases of Venus", *Journal for the History of Astronomy*, 32 (2001): 109–29.
11. The conviction that the earth shines only by reflected light led Galileo to another unforced error—the claim that the Northern Lights are caused by reflected or refracted light from the sun. For the context of this, see Eileen Reeves, *Painting the Heavens: Art and Science in the Age of Galileo* (Princeton, NJ: Princeton University Press, 1997), 57–90.
12. The key text on magnetism, of course, was William Gilbert's *De Magnete* (1600).
13. Galileo, 'On Motion', and 'On Mechanics', ed. and trans. I. E. Drabkin and Stillman Drake (Madison: University of Wisconsin Press, 1960).



14. Stillman Drake, *Cause, Experiment and Science* (Chicago: University of Chicago Press, 1981).
15. Galileo Galilei, *Two New Sciences*, trans. Stillman Drake (Madison: University of Wisconsin Press, 1974), 80-86 (abbreviated as *TNS*), OG xii 354.
16. *TNS* 11.
17. OG xiv 127-30.
18. OG xii 168.
19. *TNS* 20-6; OG xiv 127-30.
20. OG xiv 158-60. Orazio Grassi made the same error in correspondence with Baliani: Serge Moscovici, *L'expérience du mouvement: Jean-Baptiste Baliani, disciple et critique de Galilée* (Paris: Hermann, 1967), 230-63. It seems to me mistaken to claim, on the basis of this correspondence, that Grassi was engaged in a Jesuit campaign against belief in the vacuum: Pietro Redondi, *Galileo Heretic* (London: Allen Lane, 1988), 296-7.
21. I am grateful to Keith Allen for pressing me on this point.
22. Cornelis de Waard, *L'expérience barométrique, ses antécédents et ses explications* (Thouars: Imprimerie Nouvelle, 1936); William R. Shea, *Designing Experiments and Games of Chance: The Unconventional Science of Blaise Pascal* (Canton, Mass: Science History Publications, 2003).
23. Wootton, *Galileo*, 251-58.
24. Wootton, *Galileo*, 33-36.
25. Stillman Drake, *Cause, Experiment and Science*, xxv-xxix.
26. For near misses see OG iii 340 and vi 538.
27. *TNS* 110-12, 262-63 (and note 2), 272-73. On Galileo's techniques for visualising problems, Paolo Palmieri, "Mental Models in Galileo's Early Mathematization of Nature", *SHPS*, 34 (2003): 229-64.
28. Valuable on the nature of Archimedean science is Lucio Russo, *The Forgotten Revolution: How Science Was Born in 300 B.C. and Why It Had to Be Reborn* (Berlin: Springer, 2004).
29. Galileo, "On Motion", 29-30.
30. The classic text is Samuel Edgerton, Jr., "Galileo, Florentine *Disegno* and the 'Strange Spottedness' of the Moon", *Art Journal*, 44 (1984): 225-48.
31. On sunspots, the key text will be Galileo Galilei and Christoph Scheiner, *On Sunspots*, ed. Eileen Reeves and Albert van Helden (Chicago: University of Chicago Press, forthcoming 2010).
32. Galileo, "The Assayer", in *The Controversy on the Comets of 1618*, ed. Stillman Drake and C. D. O'Malley (Philadelphia: University of Philadelphia Press, 1960), 151-336, at 308-14.
33. OG iv 22, 112; Drake, *Cause, Experiment and Science*, 130, 217. Drake misquotes the key phrase and gives an incorrect page reference.
34. Steffen Ducheyne, "Galileo's Interventionist Notion of 'Cause'", *Journal of the History of Ideas*, 67 (2006): 443-64.
35. The Latin text known to Galileo reads: 'Sed causa est illud quod, quando destruetur, destruetur causatum; et quando revertitur, revertetur causatum'; or, in A. Mark Smith's modern translation: 'Now a cause is such that when it ceases to operate, what it causes ceases to exist; and when it is brought back to bear, what it causes comes back into existence.' *Alhacen's Theory of Visual Perception*, ed. A. Mark Smith (2 vols., Philadelphia: American Philosophical Society, 2001), 43-44, 370. A translation from the original Arabic, which was of course unknown to Galileo, is little different: Ibn Al-Haytham, *The Optics*, trans. A. I. Sabra (2 vols., London: Warburg Institute, 1989), i 77: 'Now a cause is such that if it ceases, the effect ceases, and if it returns, the effect returns.' For Galileo's knowledge of Alhacen, see Wootton, *Galileo*, 79.



36. David Wootton, "Accuracy and Galileo: A Case Study in Quantification and the Scientific Revolution", *The Journal of the Historical Society*, 10 (2010): 43–55, at 47.
37. Thomas S. Kuhn, "Mathematical Versus Experimental Traditions in the Development of Physical Science" [1976], in *The Essential Tension: Selected Studies in Scientific Tradition and Change* (Chicago: University of Chicago Press, 1977), 31–65.
38. Godefroy's *Dictionnaire de l'ancienne langue française* (1881) lists *esperiment*, and the *Dictionnaire du moyen français* (<http://www.atilf.fr/dmf/>) lists *expériment*.
39. John P. Wright, *The Sceptical Realism of David Hume* (Manchester: Manchester University Press, 1983), 192–209.
40. On the French Revolution, see, for example, François Furet, *Interpreting the French Revolution* (Cambridge: Cambridge University Press, 1981).



## 2 Primary and Secondary Causes in Descartes' Physics

*Tad M. Schmaltz*

In his *Principles of Philosophy*, Descartes introduces the distinction in his physics between the 'universal and primary cause' of motion, which he identifies with God, and the 'particular and secondary causes' of motion, which he identifies with 'rules or laws of nature' (*PP* II.36–37, *AT* 8–1:61–62).<sup>1</sup> This distinction derives from the scholastic tradition, and we will see that Descartes' particular account of it makes contact at several crucial points with the account of this distinction in the work of the late-sixteenth-century scholastic Francisco Suárez. But though Suárez was a critic of the occasionalist thesis that God is the only real cause, a prominent view in the literature has Descartes offering an occasionalist form of physics. I think that a closer consideration of what Descartes has to say about primary and secondary causes of body-body interactions reveals not only that he sides with Suárez in opposing an occasionalist account of the production of effects that pertain to physics, but also that he offers a more radical form of anti-occasionalism in this realm than what one finds in Suárez.

Here I begin with the main positions on causality at the beginning of the early modern period, which include not only occasionalism and Suárez's 'concurrentism', but also what Alfred Freddoso has called the 'mere conservationism' of the fourteenth-century theologian Durandus of Saint Pourçain.<sup>2</sup> All three positions agree that every finite being depends on God for its continued existence as well as for its initial creation. Whereas occasionalists hold that divine activity alone produces all effects of causal interventions, however, concurrentists and mere conservationists insist that creatures make a real though dependent contribution to such interventions. The difference between the latter two positions is that on mere conservationism divine conservation exhausts God's contribution to the natural production of effects, whereas on concurrentism there is in addition to conservation God's 'general concursus' with the actions of created agents.

After considering this background, I turn to Descartes' account of the role in his physics of God as the primary cause of motion. This account draws on the particular conception of conservation in Suárez, and also



employs Suárezian language in emphasising the importance of God’s ‘ordinary concursus’. However, I argue that the sort of concursus that Descartes invokes does not differ from divine conservation. On this point, Descartes is closer to Durandus’s mere conservationism than he is to Suárez’s concurrentism.

A mere conservationist reading of the role of God in Descartes’ physics requires that Descartes attribute real causal agency to bodies. And indeed, he does refer to particular causes of motion distinct from God. However, he offers the initially strange view that the laws of nature serve as particular causes. Some have been inclined to an occasionalist understanding of this view, on which these laws simply describe the ways in which God brings about natural changes. The motivation for this reading is clear, since Descartes’ identification of matter with mere extension seems to leave no room for the attribution of causal powers to bodies. However, I think we need to take seriously the implication of Descartes’ discussion of the laws of nature that bodies have forces that bring about changes in motion due to collision. Here I develop in a new way an anti-occasionalist account in the literature of how the bodily forces that Descartes posits could be conceived in terms of extension. I conclude that Descartes not only takes these forces to be real secondary causes, but also holds, contra Suárez’s concurrentism, that the forces, and not God’s ordinary concursus, are immediately responsible for the natural changes due to body-body interactions.

SCHOLASTIC BACKGROUND

Common to the three major Western monotheistic traditions is the view that God is the primary and provident cause of nature. At the beginning of the early modern period, there were the following three main options concerning the relation of God to nature.

OCCASIONALISM	CONCURRENTISM	MERE CONSERVATIONISM
(Mutakallimûn, Biel) God alone causes change in nature— ‘secondary causes’ serve merely as ‘occasions’ for divine action	(Thomas, Suárez) God not only creates/conserves creatures with powers but also ‘concurs’ with the action of those powers; middle way between occasionalism and mere conservationism	(Durandus) God creates/conserves creatures with causal powers, but only those powers are directly and immediately responsible for natural change

Though occasionalism was popular during the medieval period among the Islamic Mutakallimûn (primarily from the period spanning the tenth and the twelfth centuries) and received some support from the fifteenth-century theologian Gabriel Biel, it was a minority position at the beginning



of the early modern period.<sup>3</sup> The most dominant position at that time was concurrentism, according to which God acts immediately via *concursus* with created efficient causes. Thomas Aquinas defended this position in the thirteenth century,<sup>4</sup> and it was further developed in the work of later scholastics such as Suárez. But another, and in some ways more extreme, alternative to occasionalism was proposed by Durandus. In the second book of his commentary on Lombard's *Sentences*, Durandus agrees with Thomas in rejecting the claim that God is the sole immediate cause of all effects in nature.<sup>5</sup> However, he also rejects Thomas's conclusion that God acts with secondary causes on the grounds that if these causes are to have any efficacy at all, it must be the case that they alone immediately produce their effects.<sup>6</sup>

The point of contact among the three views is the requirement that God continually conserve the world in order for it to remain in existence. All the main parties agree that created objects are dependent on God for existence at all times subsequent to their creation for the same reason that they are at the instant of creation, namely, that all beings other than God are only 'beings by participation' that depend for their existence on God.

In Suárez's *Disputationes Metaphysicae*, there is the further claim—as we will see, relevant to Descartes' views—that the conservation of an object differs from the initial creation of it 'only by reason of a certain connotation' and not 'with respect to reality or real mode', that is to say, the act of creation and the act of conservation are distinct merely 'in reason' and not in reality (*DM* 21.2.3–7, 1:791f). One argument for this claim is that in the case of the creation and subsequent conservation of an object, both the act—the granting of being (*esse*) to the object—and the effect—the actual existence of the object—are the same. Suárez admits that the time of creation differs from that of conservation, but holds that since the presence of *esse* is continuous during this time, the action can be considered to be the same. The only difference between creation and conservation is that creation connotes a prior absence of existence,<sup>7</sup> whereas conservation connotes a prior existence.

These views can be accommodated to both occasionalism and mere conservationism. What would be unacceptable to partisans of these alternative theories, however, is Suárez's account of God's 'concursus' with secondary causes. Suárez attempts to hold, contrary to the occasionalist, that secondary causes have real efficacy, while holding contrary to the mere conservationist that creatures cannot produce anything without the immediate assistance of God. I set aside here Suárez's rejection of occasionalism,<sup>8</sup> and focus on his argument against mere conservationism. The stated target of this argument is Durandus. In book 2 of his commentary on the *Sentences*, Durandus grants the necessity of divine conservation for existence, but denies the view of the pre-Suárezian Thomistic concurrentists that an immediate divine concursus is required for the action of secondary causes. For Durandus, God produces the effects of secondary causes only



mediately by conserving the existence of these agents and their causal powers (*In sententias*, bk. II, dist. 1, q.5).

Suárez argues against Durandus that concurrentism follows from the conservationist thesis that all beings depend immediately on God for their continued existence. If effects did not so depend, they would have an *esse* independent of God's activity. But the doctrine of mere conservationism is that all created beings depend immediately on God's activity for their *esse*. So the effects of created causes also depend on God's activity for their *esse*, contra mere conservationism (*DM* 22.1.7–12, 1:992–94).

For Durandus, however, it is impossible that the effect depend immediately both on God and on secondary causes. Here he offers, in response to concurrentism, what we could call the 'argument from divided effect'. Either God causes the whole of the effect, or causes only part. If the whole, then the contribution of secondary causes is superfluous; we land in occasionalism. But if only part, then there is a part of the effect that derives immediately only from the secondary cause and not from God, which is Durandus's own position (*In sententias*, bk. II, dist. 1, q.5, §§11–14). This argument depends on the premise that 'it is impossible for numerically the same action to be from two or more agents in such a way that it is immediately and completely from each, unless numerically the same power is in them' (*ibid.*, §12). Since God and creatures cannot have numerically the same power, if God is the immediate and total cause of an effect, nothing else can be.<sup>9</sup>

Suárez responds with the powerful argument that there can be two total causes of the same effect if one is subordinated to the other (*DM* 22.1.22, 1:807). One example of this would be where the author and her pen are both total causes of the words on a page, with the pen being subordinated to the author. So also, Descartes seems to allow that created causes are subordinated to God when he notes in a 1645 letter to Princess Elisabeth that God is 'the universal cause of everything in such a way as to be the total cause of everything; and so nothing can happen without his will' (6 Oct. 1645, AT 4:314). Though this passage has been read as excluding 'total causes' other than God,<sup>10</sup> the passage itself emphasises only that God's universal causality differs from the universal causality of the sun insofar as the latter depends on particular causes not subordinated to it.<sup>11</sup> Thus, the point is not that there can be no derivative total causes, but only that any such causes must be subordinated to God's universal and total causality.

However, there is another argument for mere conservationism that Durandus himself rejects, but that is relevant to Descartes' views. As Durandus states it, the argument is that 'things have being insofar as they are from God, since in this they are not distinct from one another, but insofar as they have distinct being they are from secondary causes, through which they are differentiated'. Durandus judges this argument to be 'defective' on the grounds that it claims falsely 'that God acts uniformly in all things and that there is no diversity in things except from secondary causes'



(*In sententias*, bk. 11, dist. 1, q.5, §9). This objection assumes, questionably, that the claim here is the strong one that God *cannot* produce diverse effects by himself. But there is the possible position that God naturally produces only a uniform effect, with nonuniformities being attributable to secondary causes alone. This is the kind of mere conservationism that I take to inform Descartes' account of God as the primary cause of motion.

## GOD AS PRIMARY CAUSE

In article thirty-six of the second part of the *Principles*, Descartes begins his discussion of the cause of motion by distinguishing between God as 'universal and primary' or 'general cause' of motion and 'secondary and particular causes' that 'produce in an individual piece of matter some motion that it previously lacked' (AT 8-1:61). What we seem to have here is a distinction between God as general cause of the total quantity of motion (which is the sum of the quantities of motions in particular moving bodies, measured by the products of the volumes of those bodies and the speeds of their motions) and particular causes that produce changes in the distributions of that quantity.<sup>12</sup> Descartes speaks in this article of God as conserving the total quantity solely by means of his 'ordinary concursus', and thus seems to accept a version of concurrentism.<sup>13</sup> But it turns out that Suárez's account of divine conservation, rather than his account of divine concursus, is most relevant to Descartes' understanding of God's activity as the primary cause of motion.

Suárez's view of the identity of creation and conservation is broached in *PP* II.42: God conserves the world continually 'through an action identical with his original act of creation' (AT 8-1: 66). Descartes' position here is connected to his argument in the Third Meditation that it follows 'from the nature of time' that 'a work of the same force and action is needed to conserve a thing at each moment it endures as the work to create that same thing anew, if it did not yet exist; and so conservation differs solely in reason from creation, and this is one of those things that are manifest by the natural light' (AT 7: 49). Martial Gueroult and others have taken Descartes to argue here that since each moment is distinct from all other moments, God must re-create an object at each separate moment at which it exists. Call this the 'cinematic view' of divine conservation (with each separate moment of existence serving as a distinct 'frame').<sup>14</sup> But Descartes does not say that God must create an object anew at each moment of its existence; rather, his claim is that God must conserve the object at each moment with the same force and action that he *would need* to create it anew. Moreover, the conclusion that conservation is only 'distinct in reason' from creation seems clearly to be in line with Suárez's account of conservation, on which conservation does not consist in a series of distinct acts of creation, but rather is *identical* to the initial act of creation. This Suárezian position



appears in particular to be reflected in the conclusion in the *Principles* that God conserves the world through ‘the same action’ as his original act of creation (PP II.42, AT 8–1:66).<sup>15</sup>

I have noted the claim in *Principles* II.36 that God acts subsequent to creation ‘by his ordinary concursus alone’ (AT 8–1:61). In order to understand the precise meaning of this claim, let us start with *ordinary* and then move to *concursus*. There was a common scholastic distinction between God’s ‘absolute’ and ‘ordinary’ (or ‘ordained’) power. Drawing on the account of this distinction in Thomas, Suárez holds that God’s absolute power (*potentia absoluta*) is his power to affect anything ‘apart from any respect toward the nature of things or toward other causes’, whereas his ordinary power (*potentia ordinaria*) is involved in his action ‘according to the common laws and causes that he has established universally’ (DM 30.17.32, 2:216).<sup>16</sup> It is because God’s power is absolute in this way that he is able to produce by means of miracles logically possible effects that do not follow from the natures of objects in the created world. In contrast, God produces all effects that follow from these natures by means of his ordinary power. By excluding from consideration any changes in the material world guaranteed ‘by divine revelation’ (PP II.36, AT 8–1:61), Descartes indicates as well that God’s ordinary concursus excludes the miraculous actions that God can produce by his absolute power.<sup>17</sup>

But even though the scholastic background helps us to understand the sense in which the ordinary concursus that Descartes invokes is *ordinary*, there is reason to doubt that Descartes is working with the same notion of *concursus* as the one found in the work of concurrentists such as Suárez. We have seen that Suárez denies any real distinction between God’s initial creation of an object and the act by which he conserves that object over time. However, he insists on the distinction of divine conservation from the acts by which God concurs with secondary causes. Suárez argues that since ‘the concursus external to God is nothing other than the action itself’ by which the secondary cause acts, ‘the concursus will vary according to the variety of the actions’ (DM 22.4.8, 1:831). Whereas God immediately conserves an object at different times by means of the same act, then, he must concur by distinct acts in the different operations of that object.

In contrast, Descartes consistently holds that there is only a single unvarying action that is involved in the ordinary concursus that constitutes God’s contribution to the causation of motion. Just as he claimed in *Le Monde* (c. 1633) that the laws of nature ‘depend only on the fact that God conserves each thing by a continued action’ (LM VII, AT 11:37), so he claims later in the *Principles* (1644) that the third law of nature, which governs the communication of motion (see the discussion of this law following), ‘is demonstrated by the immutability of the action of God, continually now conserving the world by the same action by which he created it then’ (PP II.42, AT 8–1:66).



In a response to my claim in another work that Descartes does not distinguish God's ordinary concursus in physics from his conservation of the material world, John Carriero has objected that in *Principles* II.36, at least, 'Descartes does not indicate . . . that God's ordinary concursus is limited to his conservation of the total quantity of motion and rest. All he says is that the conservation of motion falls under God's ordinary concursus, so that a motion's continuing indefinitely is not miraculous'.<sup>18</sup> But at least with respect to the purely material world, Descartes defines God's ordinary concursus in terms of the fact that he 'now conserves the whole of this matter in the same way and with the same plan [*ratione*] by which he first created it' (AT 8-1:62). Moreover, if by the point that 'motion's continuing indefinitely is not miraculous' Carriero is referring to the inertial tendency of a particular moving body, this concerns Descartes' first law of nature (to be discussed following) rather than the principle of the conservation of the total quantity of motion that is at issue in this portion of the *Principles*.<sup>19</sup> The point in this portion is rather that the fact that 'the same motion is always conserved' in the material world as a whole follows immediately from the nature of God's ordinary concursus (AT 8-1:62).

Carriero has further objected that my reading of Descartes is 'puzzling' since it seems to preclude the claim that God's ordinary concursus is responsible for varying effects. The problem here is that this implication makes it difficult for Descartes to 'find some way of reconciling God's immutability with the fact that God's causation is involved in a world that changes'.<sup>20</sup> Suárez's way around the problem of reconciling a changing concursus with divine immutability is to claim that what changes is only the 'external concursus' that resides in the patient, and not 'the internal concursus or volition by which God concurs' (DM 22.4.8, 1:831). However, this line of response cannot be sufficient for Descartes, since he is clear not only that the principle of action in God is immutable, but also that what God produces is always the same. Thus, Descartes notes in *Le Monde* that it follows from divine immutability that God 'always produces the same effect', and in particular always produces 'a certain quantity of motions in matter in general' (LM VII, AT 11:43). And though Descartes allows in the *Principles* that God could produce changes by means of acts that go beyond his ordinary concursus, he continues to hold that this concursus itself yields a constant effect.

To be sure, this line of argument still may seem to be puzzling. After all, Descartes granted, with Suárez, that an eternal principle in God can yield temporal effects. So why not allow, with Suárez again, that an immutable divine principle can yield an inconstant concursus? Here I think it is best to understand Descartes' thought in terms of Suárez's account not of the divine concursus, but rather of divine *conservatio*. Suárez argues that God's creation of an object cannot be distinguished from his subsequent conservation of it since in both cases the agent is producing in the same



way, namely, *ex nihilo*, the very same effect, namely, the *esse* of the object. We can derive the sameness of the act from the fact that both the production and the terminus remain the same (DM 21.2.3, 1:791). Since Descartes assumes that God produces motion as a primary cause by means of a single concursus identical to his continuing creation of matter in motion, he needs to emphasise that the effect of this divine act remains the same.

We do not have to speculate that Descartes' account of divine concursus would have been objectionable to those with a Suárezian sensibility. One of Descartes' Dutch critics, Jacobus Revius, the author of *Suarez repurgatus*,<sup>21</sup> takes issue this account when he announced in a 1650 text that

I am ashamed for the love of God, I am ashamed about your ignorance, Descartes. That such a great philosopher as yourself has not learned to distinguish between *conservation* and *concursus*!<sup>22</sup>

Revius makes what for scholastics would be the crucial point that an object that is nothing apart the divine act of conservation cannot be said to concur in that act.<sup>23</sup> In the scholastic view, creatures concur not in the conservation of their existence, which God alone can bring about,<sup>24</sup> but only in the production of their effects, to which their operations contribute.

Revius also questions how Descartes could speak of concursus at all in the context of his physics given that there can be no *vis agendi* in his world of mere extension, and thus no 'acting with' on the part of bodies.<sup>25</sup> Such a question, of course, broaches the issue of Descartes' views on the efficacy of secondary causes in his physics, an issue that it is now time to confront directly.

## LAWS AS SECONDARY CAUSES

Descartes' claim in the *Principles* that 'rules or laws of nature' (*regulae sive leges naturae*) are 'particular and secondary causes' (PP II.37, AT 8–1:62) no doubt strikes us as somewhat odd. From our post-Humean perspective such rules or laws would seem to be mere empirical generalisations, hardly the sort of entity that could serve as a cause. In order to discern what Descartes could have had in mind in calling his laws causes, however, we need to attend to his particular description of the laws.

I focus here on the first and third of his three laws of nature, which as we will see are most directly relevant to the production of changes in the distribution of the quantity of motion among particular bodies.<sup>26</sup> The two laws are stated as follows:

1. Each and every thing [*res*], insofar as it is simple and undivided, always perseveres in the same state, insofar as in it lies [*quantum in se est*], and never changes except by external causes. (PP II.37, AT 8–1:62)



3. Where a body that moves encounters another, if it has a less force for proceeding [*vim pergendum*] in a straight line than the other has for resisting it [*resistendum*], then it is deflected in another direction [*in aliam partem*], and retaining its motion gives up only the determination of motion; but if it has more, then it moves the other body with it, and the quantity of its motion it gives to the other it also loses. (*PP* II.40, AT 8–1:65)

In *Principles* II.36 the 'particular cause' is said to be that 'which makes it so that singular parts of matter acquire motion they did not have previously' (AT 8–1:61), and of the two laws just cited it is the latter that provides the best example of a particular cause in this sense. According to the third law, what does the causal work in bringing it about that parts of matter acquire a new quantity of motion is the 'force for persevering' that produces a communication of motion.<sup>27</sup>

Even so, Descartes indicates that this third law depends in an essential way on his first law. Thus he claims in *Principles* II.43 that 'the force of any given body to act on another' consists 'in this one thing, that each and every thing tends, *quantum in se est*, to persist in the same state it is in, as posited in the first law' (AT 8–1:66). It is in virtue of the fact that a moving body has a tendency to persist in the same state *quantum in se est* that it has the force to persist in its motion. The indication here seems to be that the efficacy of the third law as a particular cause is to be explained in terms of internal bodily tendencies that are themselves responsible for changes in the distribution of the quantity of motion.

To be sure, we have seen Descartes' claim in the *Principles* that the third law is due to the immutability of God's conservation of the world. It is certainly possible to read Descartes' talk of inherent forces and tendencies—as Daniel Garber, for instance, has recently—as a mere *façon de parler*, a way of describing the changes that God alone produces by means of his continuing conservation of the material world.<sup>28</sup> And there seems to be a clear motivation for this sort of occasionalist reading. For on Descartes' official position in the *Principles*, 'extension in length, breadth and depth constitutes the nature of corporeal substance', and thus 'everything else that can be attributed to body presupposes extension, and is merely a mode of an extended thing' (*PP* I.53, AT 8–1:25). It appears that whatever forces for persisting are, they cannot be mere modes of extension akin to various shapes and motions. But then it seems that such forces would be similar to the 'real qualities' that the scholastics posited as secondary causes in nature and that Descartes himself wanted to banish from physics. Surely Descartes would have wanted to banish occult bodily forces as well?

Before acquiescing to the occasionalist interpretation, it is worth pausing to consider more carefully the comparison to scholastic real qualities. In the case of Suárez, there is in fact a complex theory of the causal efficacy of accidental features of bodies. This theory starts from Aristotle's list of



predicables, which distinguishes the basic accidental categories of quantity, quality, relation, action, passion, time, place, position, and having. Of these categories, Suárez holds that only certain qualities, and neither quantities nor relations, can be principles of action. In particular, he explicitly denies that either shapes (in the category of quality)<sup>29</sup> or local motions (as well as alteration in quality, augmentation in quantity, and substantial generation) can serve as *per se* principles of action (see *DM* 18.4, 1:624–27).

Clearly, Descartes rejects the assumption in Suárez that there are principles of action in bodies other than quantity, shape, and local motion. But does he accept the Suárezian position that these other features of bodies cannot be principles of action? If so, Descartes would be forced to conclude that there are no such principles in bodies. However, it seems to me that the claim that Descartes in fact draws this conclusion is often based less on the textual evidence than on intuitions about what his identification of body with extension requires. In particular, the guiding intuition is that bare extension is something that is purely passive, the mere instantiation of a purely geometrical essence, and not something that can ground causal activity.

Admittedly, some of Descartes' statements may seem to indicate that he conceived the extension that constitutes the nature of bodies that actually exist in the material world in just this manner. There is, for instance, his claim in the *Principles* that 'I admit no other principles in physics but those in geometry or abstract mathematics' (*PP* II.64, AT 8–1:78), as well as his earlier remark to Mersenne that 'my physics is nothing but geometry' (27 July 1638, AT 2:268). But Descartes' own account of the laws of nature belies this simple identification of physics with geometry. For an object of geometry, as present objectively in the mind that considers it, does not have the tendency to persist in a particular state *quantum in se est* (as required by the first law), and does not have any force that in cases of collision requires changes in the states of other bodies (as dictated by the third law). The reason that purely geometrical objects lack these features seems to be that they lack any sort of existence external to mind. But could it be that they possess the features simply in virtue of possessing this additional extramental existence?

The answer to this question is 'yes', according to the most sophisticated argument in the literature for the view that Descartes attributes real causal efficacy to bodies. Thus, Alan Gabbey has urged that Descartes distinguishes 'between a body's essence as an idea, that is as existing objectively in the intellect, and the body's existence outside the mind', and that 'force depends on extension in the sense that extension is presupposed in saying that something corporeal exists or endures'.<sup>30</sup> The difference between a purely geometrical object and an actually existing body is that only the latter possesses an existence or duration that, on Descartes' official view in the *Principles*, is distinct only by reason (*rationis*) and not in reality (*in re*) from the substance that exists and endures (see *PP* I.62, AT 8–1:30). If



force is simply identified with the attribute of existence or duration, then it seems to be something that concretely existing bodies, but not abstract geometrical objects, possess.

There is the question, however, whether the various forces for persisting that Descartes posits in the *Principles* can be explained in terms of the forces that Gabbey identifies with the durational existence of bodies. It seems that they cannot be so explained, for the simple reason that whereas Descartes offers existence or duration as an example of something 'which always remains unmodified', and so is an attribute in the strict sense (*PP* I.56, AT 8-1:26), his view that bodily forces constantly change due to collision indicates that such forces are variable, and so cannot be attributes strictly speaking.<sup>31</sup>

Gabbey anticipates this problem, however, and responds to it by appealing to the fact that there is in addition to the attribute of duration a force that is present in bodies as a variable mode rather than as an attribute. He explains this additional force in terms of the scholastic distinction—which Descartes endorses in his response to Gassendi in the Fifth Replies—between causes of being (*secundum esse*) and causes of becoming (*secundum fieri*) (see AT 7:369). On Gabbey's reading of Descartes, 'force as *causa secundum esse* is . . . an attribute of body, in the sense that *qua* cause it is necessarily entailed by a body's duration, viewed *simpliciter* and irrespective of mode', whereas when the forces are viewed 'as quantifiable causes of change in the corporeal world, or as reasons . . . explaining absence of change of a certain kind in particular instances, they are *causae secundum fieri*' that are 'clearly in body *diverso modo*, and so are modes of body, rather than attributes'.<sup>32</sup>

But when forces are present in bodies *diverso modo*, what sort of modes are they? Gabbey does not say explicitly, but there is an answer that draws on the implication of Descartes' first law that duration involves a tendency to persist in the same state. An initial suggestion is that these tendencies be conceived as variable modal features of duration. Given this conception, we could perhaps identify the varying forces that serve as causes *secundum fieri* with such features.

Gabbey himself rejects this suggestion on the grounds that 'a body cannot have "more or less" duration: either it exists or it does not, and if it does exist, in whatever modal disposition, it necessarily endures without "existential variation"'.<sup>33</sup> As I have indicated, Descartes does indeed say in the *Principles* that duration is an unvarying attribute rather than a varying mode. However, he indicates that thought and extension also are attributes that are not subject to modification *considered as such*. As he notes in correspondence with Arnauld,

as extension, which constitutes the nature of body, differs greatly from various shapes or modes of extension that it assumes, so thought, or thinking nature, in which I take the essence of the human mind to



consist, is much other than this or that act of thinking. (29 July 1648, AT 5:221)

In the same way, it seems, Descartes could say that invariable duration differs greatly from the various modifications it assumes.

My proposal that Descartes' bodily forces are modal features of bodily duration is still rather abstract and somewhat disconnected from his discussion in the *Principles* of the various *vires ad pergendum* involved in the communication of motion.<sup>34</sup> In order to make the proposal more concrete, I start with the point in that text at which the third law of nature makes contact with the first. I have noted the claim in *Principles* II.43 that the forces that the third law takes to be responsible for changes in motion due to collision consist simply 'in this one thing, that each and every thing tends, *quantum in se est*, to persist in the same state it is in, as posited in the first law' (AT 8-1:66). What is *in se*, according to the first law, is the tendency to continue in the same state. What the later use of this law makes clear is that the strength of this tendency is not constant but varies depending on the nature of the mode involved. The measure of the strength of the tendency to persist in motion is just the same as the measure of the forces involved in this state. Thus, in the case where one body is double the size of another body moving at the same speed, the first body has double the strength of the tendency to persist in its state of motion than the other body has to persist in its state. The tendency can be said to be something in what has the tendency insofar as it is simply a modal feature of the duration of a moving body.

We have seen the view in the *Principles* that duration as an attribute is distinct only *rationis*, and not *in re*, from the substance that endures (AT 8-1:30). On the view I attribute to him, though, Descartes also holds that the strength of the tendency to endure in a state of motion, or force of proceeding, is not distinct in reality from the state that possesses a tendency with this strength. One reason that it seems that force is not intrinsic to these modes is that the modes themselves can be considered abstractly as purely mathematical objects, the nature of which is exhausted by their geometric and kinematic aspects. In this respect, these modes are similar to the triangle that, as Descartes tells an unknown correspondent, can be considered merely with respect to its essence, in abstraction from its existence. In this case, according to Descartes, the thought of the essence of a triangle differs modally from the thought of the existence of that triangle. But he continues by noting that outside of thought the essence of a triangle and its existence are 'in no way distinct' (1645 or 1646, AT 4:350). Though Descartes could allow similarly that the thought of force, that is, of the strength of the tendency to endure, is modally distinct from the thought of the modal features of bodies that have that force, still it seems open to him to say that the force or strength of tendency to endure is in no way distinct from those features as they exist external to mind.



Given this distinction between the two ways of considering modes of motion, we can see that there is an important ambiguity in Descartes' official position that all bodily modes must be conceived through extension. For as Gabbey has emphasised, extension itself can be conceived merely abstractly, as present objectively in the intellect, or concretely, as something not distinct from its durational existence in reality. If the modes are conceived in terms of abstract extension, then force is not an intrinsic feature of them. Their nature would be exhausted by their purely mathematical features. But if they are conceived in terms of concrete extension, force can be intrinsic to them insofar as this force is identified with the strength of the tendency to endure that does not differ in reality from the enduring modes.

Now let us consider how, on Descartes' view, God's activity as primary cause is supposed to be related to the activity of bodily forces as secondary causes. The indication in the *Principles* is that God's ordinary concursus consists simply in the fact that he 'diversely moved parts of matter when he first created it, and now conserves the whole of this matter in the same way and with the same plan by which he first created it' (PP II.36, AT 8-1:61). Given my reading earlier, we can take the claim here to be that God originally created moving parts of matter with tendencies to endure that have quantifiable degrees of strength. Simply by conserving the total quantity of the durational strength of moving bodies, God conserves the total quantity of motive force.

If this reading is correct, there can be in Descartes' physics no difference between God's ordinary concursus and his continuing conservation of the total quantity of matter and its motion. Since this continuing action must have a constant effect, the changes in motion produced by bodily collisions must be due not to that action, but rather—as Descartes himself indicates—to the secondary causes of motion. What we have here is not the position in Suárez that God concurs by means of an action distinct from his act of conservation. Instead, Descartes is closer to the view of Suárez's opponent Durandus that though God is the immediate cause of the being of secondary causes, he acts only mediately in the action of such causes through his conservation of them. In Descartes' case, the view is that God's ordinary concursus is exhausted by his continuing conservation of matter with the total quantity of the force of its parts. The force of these parts, rather than the divine concursus, is the immediate cause of changes due to bodily collisions.

Far from placing Descartes with the occasionalists, then, I see him as advocating in the *Principles* an account of causation in physics that is even less concessive to occasionalism than is Suárez's concurrentism. On my reading of this text, Descartes is simply endorsing a line of thought present in his earlier work, *Le Monde*. In the latter text, for instance, he notes that

by *Nature* I do not understand some deity, or some other force of imaginary power, but I use this to signify Matter itself, being that I consider



it with all the qualities that I have attributed to it, and under the condition that God continues to conserve it in the same manner that he has created it. Because from that alone, that he continues to conserve it, it follows, of necessity, that there must be several changes in its parts, which cannot, it seems to me, be properly attributed to the action of God, because it never changes. (LM VII, AT 11:37)

I take the view here that changes in parts must be attributed to ‘matter itself’ rather than to the action of God to be reflected in the later implication in the *Principles* that bodily forces, and not God’s ordinary concursus, serve to explain alterations in the distributions of the quantity of motion due to collision.<sup>35</sup> If I am correct, Descartes’ most considered position is that these forces include forces of moving bodies that are to be explained in terms of the strengths of their tendencies to endure. Only these forces can explain changes in motion, since divine conservation always produces the same effect, namely, the same total quantity of motive force, that is, strength of tendency to endure in motion. The suggestion here is that God’s ‘immediate concursus’ involves only his conservation of this total quantity, and not his production of the particular effects attributable to matter. Though he may have spoken with Suárez, the evidence seems to me to indicate that Descartes consistently thought with Durandus.

## NOTES

1. Abbreviations used: AT: C. Adam and P. Tannery (eds.) *Œuvres de Descartes*, 11 vols. (Paris: Vrin, 1964–74). DM: Francisco Suárez, *Disputationes Metaphysicae*, 2 vols. (reprinted, Hildesheim: Georg Olms, 1965); cited by disputation, section, paragraph. For an English translation of relevant portions of this text, see A. Freddoso (trans.) *On Creation, Conservation and Concurrence: Metaphysical Disputations*, 20–22 (South Bend, IN: St. Augustine’s Press, 2002). *In sententias: In sententias theologicas Petri Lombardi* (Venice, 1571; reprinted, 2 vols., Ridgewood, NJ: Gregg Press, 1964). For an English translation of bk. II, disp. 1, q. 5 of *In sententias*, see A. Freddoso’s entry on <http://www.nd.edu/~afreddos/translat/duran215.htm>. LM: *Le Monde*, cited by chapter. PP: *Principles of Philosophy*, cited by part and article. ST: Thomas Aquinas, *Summa Theologica*; for an English translation of relevant portions of this text, see *The Basic Writings of Saint Thomas Aquinas*, vol. I, ed. A. Pegis (Indianapolis: Hackett, 1997).
2. See Alfred Freddoso, “God’s General Concurrence with Secondary Causes: Why Conservation Is Not Enough”, *Philosophical Perspectives*, 5 (1991): 553–85.
3. For a general study of medieval and early modern occasionalism, see Dominik Perler and Ulrich Rudolph, *Occasionalismus: Theorien der Kausalität im arabisch-islamischen und im europäischen Denken* (Göttingen: Vandenhoeck & Ruprecht, 2000).
4. For instance, in ST Ia.105, a.5.
5. ‘[E]ven though a few a few thinkers at one time maintained that creatures do nothing at all (e.g., that fire does not produce heat, but that instead God



creates the heat, and so on for all other creatures), nonetheless, this view is now rejected by everyone as implausible, because it denies of things their proper operations and also denies the sensory judgment by which we experience that created things act on one another' (*In sententias*, bk. II, dist. 1, q.5, §4).

6. In this *questio*, Durandus cites *ST* Ia.103, a.6 and 105, a.5.
7. Suárez allows for creation from eternity (*ab aeterno*), which, in contrast to the case of creation in time (*de novo*), does not presuppose any temporally prior absence of existence. Even so, creation *ab aeterno* is said to presuppose a 'metaphysical' absence of existence that consists in the fact that the created thing does not exist by its own nature.
8. See the position in *DM* 18.1.6, 1:594–95, which is similar to the position in Durandus cited in note 5.
9. Durandus also thinks that the concurrentist account compromises divine goodness insofar as it makes God the author of sinful actions; see *In sententias*, bk. II, dist. 37, q.1. Suárez responds that whereas God offers only a single concursus in the case of natural agents, and so determines a particular effect, in the case of free agents he offers a concursus to alternative sorts of actions, and so makes it the case that various options are open to the agent (*DM* 22.4.14–23, 1:832–34). For discussion of Suárez's response, see Alfred Freddoso, "Suarez on God's Causal Involvement in Sinful Acts", in *The Problem of Evil in Early Modern Philosophy*, edited by E. Kremer and M. Latzer (Toronto: University of Toronto Press, 2001), 10–34.
10. I find this suggestion, for instance, in Jean-Luc Marion, *Sur la théologie blanche de Descartes: analogie, création des vérités éternelles et fondement* (Paris: Presses universitaires de France, 1991), 288–89.
11. See Descartes' claim in this letter that 'what makes the sun, for instance, the cause of flowers is not the cause of the fact that tulips differ from roses, [since] their production depends also on particular causes to which they are not subordinated' (AT 4:314).
12. See the claim in *PP* II.36 that 'when one part of matter moves twice as fast as another, and the latter is twice as large as the former, we must consider that there is as much motion in the latter as in the former (AT 8–1:61). In this section Descartes also mentions that God conserves a total 'quantity of rest', which is related to his (problematic) view that resting bodies have a force to remain at rest. Here I set aside issues involving quantity of rest and force of rest, which I discuss in *Descartes on Causation* (New York, Oxford: Oxford University Press, 2008), 96–99, 120, 124, 127–28, 176. See also note 34.
13. There has been some attempt recently to read Descartes as endorsing a version of concurrentism. See, for instance, Andrew Pessin, "Descartes' Nomic Concurrentism: Finite Causation and Divine Concurrence", *Journal of the History of Philosophy*, 41 (2003), 1: 25–50, and Helen Hattab, "Concurrence or Divergence? Reconciling Descartes' Physics with his Metaphysics", *Journal of the History of Philosophy*, 45 (2007), 1: 49–78.
14. Here I am following the terminology Daniel Garber employs in *Descartes' Metaphysical Physics* (Chicago: University of Chicago Press, 1992), 266–73. Henri Bergson had previously labelled this the 'cinematographic view'; see his *L'évolution créatrice* (Paris: Alcan, 1907). For Gueroult's interpretation of Descartes' account of time, see *Descartes selon l'ordre des raisons* (Paris: Aubier-Montaigne, 1953), Vol. I, pp. 272–85. There is a more recent defence of this interpretation in Ken Levy, "Is Descartes a Temporal Atomist?", *British Journal for the History of Philosophy*, 13 (2005), 4: 627–74. Garber's weaker conclusion, which nonetheless still is contrary to what I argue here, is that Descartes' remarks, at least in the *Principles* and earlier texts, do not commit him to a rejection of the cinematic view.



15. For further argument that Descartes rejects the cinematic view of motion, see Richard Arthur, "Continuous Creation, Continuous Time: A Refutation of the Alleged Discontinuity of Cartesian Time", *Journal of the History of Philosophy*, 26 (1988), 3: 349–75.
16. Suárez cited here Thomas's discussion in *ST* Ia.25, a.5 ad.1.
17. The further exclusion of changes guaranteed 'by our own experience' most likely indicated a bracketing of changes in motions produced by finite minds. See, for instance, Descartes' claim to Arnauld that the fact that an incorporeal mind can set body in motion is revealed 'by the most certain and most evident experience [*certissima & evidentissima experientia*]' (29 July 1648, AT 5:222).
18. John Carrierio, Review of Tad Schmaltz, *Descartes on Causation*. *Notre Dame Philosophy Reviews* (<http://ndpr.nd.edu/review.cfm?id=13043>).
19. The conservation principle seems to allow, for instance, that there are changes in the quantities of motion in bodies that do not have any external cause, but that are balanced by corresponding changes in other bodies.
20. Carrierio, Review of Schmaltz.
21. *Suarez repurgatus, sive, Syllabus Disputationum metaphysicarum Francisci Suarez Societatis Jesu theologi* (Leiden, 1664). In this text, Revius attempts to present a Suárezian system purged of what to his Calvinist sensibility was Suárez's own 'Pelagian' emphasis on our freedom to reject divine grace.
22. From Revius's *Statera philosophiae cartesianae qua principiorum ejus falsitas, & domatum impuritas expenditur ac castigatur . . .* (Leiden, 1650), cited in J. A. van Ruler, *The Crisis of Causality: Voetius and Descartes on God, Nature and Change* (Leiden: E.J. Brill, 1995), 276.
23. See the passage from Revius's *Statera*, 73, cited in Van Ruler, *The Crisis of Causality*, 277, n.51.
24. Suárez allows that certain accidents can depend on created causes for their conservation (as light so depends on the sun, or motion on a mover), but insists that God alone can immediately conserve a substantial sort of being that can be created only *ex nihilo*; see *DM* 21.3, 1:794–801.
25. Again from the *Statera*, 73, cited in Van Ruler, *The Crisis of Causality*, 277.
26. The second law, which dictates that unimpeded motion is rectilinear, broaches different issues, and so will be set aside here (cf. note 27). For a discussion of this law and its implications, see my *Descartes on Causation*, 108–10.
27. Since my focus is on changes in the distributions of the quantity of motion, I will not discuss the implication of the third law that the force for resisting produces a change merely in the determination of motion (see also note 34).
28. See Garber, *Descartes' Metaphysical Physics*, 298. Cf. Gary Hatfield, "Force (God) in Descartes' Physics", *Studies in History and Philosophy of Science*, 10 (1979) 2: 113–40.
29. In opposition to a Thomistic view more common among the scholastics, however, Suárez holds that shape is a mode of quantity rather than a *res* distinct from it. This Suárezian view informs the understanding of the scholastic position reflected in Descartes' claim in a letter to Mersenne that he does not 'attribute any more reality to motion, or to all these other variations of substance that one calls *qualities*, than the philosophers commonly attribute to shape, which they call not a *qualitas realis* but only a *modus*' (26 Apr. 1643, AT 3:649).
30. Alan Gabbey, "Force and Inertia in the Seventeenth Century: Descartes and Newton", in *Descartes: Philosophy, Mathematics and Physics*, edited by S. Gaukroger (Sussex: Harvester Press, 1980), 238. Gabbey is developing the interpretation of Descartes' account of bodily force found in Martial



Gueroult, "The Metaphysics and Physics of Force in Descartes", in *Descartes: Philosophy, Mathematics and Physics*, 196–229.

31. Descartes distinguishes variable modes from invariable attributes in *PP* I.56, AT 8–1:26. The fact that the forces that Descartes invokes in his physics must be variable modes rather than invariable attributes creates particular problems for Gueroult's interpretation of Descartes, since he is clear that Descartes identifies a body's force with its duration conceived as an attribute ("Metaphysics and Physics of Force", 197). I discuss presently Gabbey's attempt to address this problem.
32. Gabbey, "Force and Inertia", 236, 237.
33. Gabbey, "Force and Inertia", 237.
34. I set aside the more problematic case of *vires ad resistendum*. The basic problem in this case, very briefly, is that whereas *vires ad pergendum* can be conceived in terms of a quantity that God continuously conserves, *vires ad resistendum* cannot be so conceived insofar as Descartes' rules of collision require that the latter appear only at contact and then disappear.
35. This is contrary to the view that Peter Machamer and J. E. McGuire offer in *Descartes' Changing Mind* (Princeton: Princeton University Press, 2009), 33–34, according to which there is a shift in the *Principles* away from the sort of realism about bodily causes found in *Le Monde*. I find the appeal in the *Principles* to bodily forces to explain changes in motion to be in line with the claim in *Le Monde* that such changes are to be attributed to features of 'matter itself'.



### 3 Causation and the Cartesian Reduction of Motion

#### God's Role in Grinding the Gears

*William Eaton and Robert Higginson*

Louis de la Forge's occasionalist explanation of local motion provides a way to reduce motion to other modes of extension. Such a reduction is favourable to the Cartesian scientist both in terms of parsimony, by reducing the total number of primary modes of extension, and in terms of intelligibility, by explaining the nature of motion at a deeper level than rival explanations.<sup>1</sup> Furthermore, it suggests an approach to the problem of reconciling problematic but important elements of Cartesian metaphysics and science. We will argue that what makes La Forge's reduction of motion possible is a commitment to solid Cartesian principles that, when taken as a whole, imply occasionalism, at least regarding body-body causation. We hold that this is both true and significant for Cartesianism in general, whether or not Descartes himself was aware of the entailment. So the first part of this chapter involves the reduction of local motion to other modes of extension, as well as the significance of this reduction, while the second part explores how the solid Cartesian principles required for this reduction entail occasionalism. In doing this we are, in part, defending the position of Daniel Garber against recent interpretations of Descartes advanced by scholars such as Helen Hattab and Tad Schmaltz.<sup>2</sup> In the end, we argue that despite significant evidence to the contrary, the entailment of occasionalism from solid Cartesian principles is so straightforward that it's hard to see how Descartes could not have been an occasionalist, at least regarding body-body causation. Our view is that occasionalism is the best way for a Cartesian to reconcile the transfer of power necessary for genuine causation with the essentially passive nature of matter.

Throughout the chapter we shall, for convenience, use the expression *solid Cartesian principle* (SCP). By *solid Cartesian principle*, we mean a principle that Descartes himself would be reluctant to give up. Solid Cartesian principles are supported by key primary sources, such as the *Meditations on First Philosophy* and the *Principles of Philosophy*, as well as



important correspondence. Solid Cartesian principles are also usually taken seriously by other Cartesians, and even when they are rejected they are treated as significant issues. Examples of solid Cartesian principles relevant to our project include the views that extension in space is the essential attribute of matter, that matter is fundamentally passive, that genuine causation involves a transfer of power, and that in God there is no distinction between volition and creation.<sup>3</sup>

Naturally, in order to understand the significance of occasionalism in the context of the Cartesian-motion-reduction project, it is necessary to know what occasionalism is all about. Occasionalism is a statement about the true nature of causality and of God's relationship to his creatures. In the seventeenth century there were a number of different versions of occasionalism, but for our purposes it will suffice to offer the following generic thesis of the view: God is the complete, exclusive, and proximate cause of anything that requires a cause.<sup>4</sup> In other words, only God is causally efficacious. This is the strongest statement of occasionalism, and it was certainly defended by Géraud de Cordemoy and Nicolas Malebranche. Other occasionalists, such as Louis de la Forge, held that bodies are completely causally inefficacious but finite minds do have their own independent power to produce certain effects.<sup>5</sup>

One of the goals of this chapter is to show that some forms of occasionalism, employing only solid Cartesian principles, allow a beneficial reduction of motion that has significant explanatory power. La Forge employs such principles in his occasionalist explanation of motion. In the *Treatise on the Human Mind* (1664), he writes:

I also claim that there is no creature, spiritual or corporeal, which can cause change in it or in any of its parts, in the second moment of their creation, if the creator does not do so himself. Since it was He who produced this part of matter in place *A*, for example, not only must he continue to produce it if he wishes it to continue to exist but also, since he cannot create it everywhere or nowhere, he must put it in place *B* himself if he wishes it to be there. For if he put it anywhere else there is no force capable of removing it from that location (FHM 147/FOP 240).

Since God must conserve every object at each moment of its existence, he must always place the object in some location or other. Motion is nothing more than the conservation of the object over time at different locations. Thus, on this account local motion turns out to be a sort of divine teleportation in which an object is *re-created* at successive locations. The object never ceases to be, but it continues its existence at different locations. In recent scholarship, this theory has come to be called, often with contempt, the *Cinematic View of Motion*. No doubt, La Forge was influenced by a now famous passage from the *Third Meditation*:



For a lifespan can be divided into countless parts, each completely independent of the others, so that it does not follow from the fact that I existed a little while ago that I must exist now, unless there is some cause which as it were creates me afresh at this moment—that is, which preserves me. For it is quite clear to anyone who attentively considers the nature of time that the same power and action are needed to preserve anything at each individual moment of its duration as would be required to create that thing anew if it were not yet in existence. Hence the distinction between preservation and creation is only a conceptual one, and this is one of the things that are evident by the natural light (CSM II, 33/AT VII, 49).<sup>6</sup>

This passage seems to claim that the preservation of an object requires a separate creative act from God for each instance of its existence. However, it should be noted that recently Tad Schmaltz has argued that such conservation need not involve the literal continued re-creation suggested in the preceding passage. According to Schmaltz, the cinematic view cannot be reconciled with Descartes' claim in the *Principles of Philosophy* that when God initially created bodies he also 'moved the parts of matter diversely, when he first created them' (CSM II, 240/AT VIII A, 62). If the cinematic view were correct, and motion is nothing more than the successive occupation of different locations at different times, then God requires more than one instant to give a body motion, so he couldn't have given it at its initial creation.<sup>7</sup>

What God gives bodies in the creation, on Schmaltz's interpretation, is a tendency or dispositional force to stay in the same state (motion or rest).<sup>8</sup> Schmaltz claims that Descartes should not be taken literally when he uses the term *re-creation*. Rather, God's conservation of objects is a continuation of his initial creative act. Schmaltz rejects the cinematic view's temporal atomism as well as its postulation of separate divine acts for each event. He feels that these features violate the simplicity and immutability of God's nature.

But we feel that the cinematic view is neither committed to separate creative acts nor literal re-creation. God's continued conservation alone, at different locations over time, is enough to make the reduction of motion work. The problem with Schmaltz's interpretation is that it fails to satisfy an important SCP, namely the passivity of matter. As Walter Ott points out in the intriguing *Causation and Laws of Nature in Early Modern Philosophy*, the tendency to remain in the same state is not a force or property of bodies, but simply God's volitional conservation.<sup>9</sup> Thus the cinematic view provides a way to mechanically explain this tendency. Furthermore, God's conservation of an object's existence at different locations alone is sufficient to explain motion. Any appeal to internal forces, even as dispositions or tendencies of the object, is superfluous. That is, once you grant such conservation, there is nothing left about motion to be explained. Consider



the matter this way: God could not conserve a body's existence at different locations over time without that body moving.

Schmaltz also claims that the cinematic view is committed to temporal atomism. Our interpretation is that the passage in the *Third Meditation*, La Forge's motivation for the cinematic view, suggests that time, like space, is indefinitely divisible;<sup>10</sup> *for a lifespan can be divided into countless parts*. However, our finite minds can only focus on specific moments. But in the case of motion, God, at some point, ceases to conserve the existence of matter at one location and starts to conserve it at another. Thus La Forge realised that such conservation could provide a straightforward way to explain the nature of local motion. God creates each body at the place he wants it to be at that moment. The change of location at each point of creation (or conservation), over time, is sufficient to produce and ultimately explain motion. This is significant to Cartesian science because it allows us to reduce motion to other modes of extension. Local motion is reducible to location plus duration via divine conservation.

Such reductions are important and not unheard of in Cartesian science and the larger mechanical philosophy. For example, consider the mechanical affections of matter advanced by Robert Boyle, which play a role in Boyle's philosophy very similar to Descartes' modes of extension. Boyle recognised some eleven total mechanical affections of matter, but they are ultimately reducible to just size, shape, motion, and the contrivance of parts. Texture, for instance, is very important in many of Boyle's explanations of natural phenomena, including fluidity and cohesion just to name a few, but it is ultimately reducible to size, shape, and contrivance of the outlying corpuscles of a corpuscular aggregate.<sup>11</sup>

A similar reduction is possible with La Forge's account of motion. This reduction is particularly significant given the importance of motion to Cartesian science. Motion is by far the most important mode of extension, and does the bulk of the explanatory work in Cartesian science. All three primary laws of nature in the second part of the *Principles of Philosophy* are laws of motion.<sup>12</sup> Furthermore, Descartes' explanations of phenomena such as heat, cold, fluidity, solidity, refraction, light, electricity, the circulation of blood, and other functions of the human body are explained in terms of motion. Finally, in the *Principles of Philosophy*, Descartes goes so far as to claim that all the variety and diversity we find in the forms of matter depend on motion (See CSM I, 232/AT VIII A, 52). It is therefore significant if motion itself is reducible to other, more basic modes of extension.

The Cartesian reduction of motion is important in terms of parsimony. Size and shape seem to be basic modes of extension, neither reducible to the other. Most mechanical philosophers accepted motion as a similar basic mode, irreducible to anything else. In *About the Excellency and Grounds of the Mechanical Hypothesis*, Boyle goes so far as to claim that '[there cannot] be any physical principles more simple than matter and motion; neither of them being resolvable into any thing'.<sup>13</sup> La Forge's explanation of motion



allows us to see that motion is actually a compound mode of extension, similar to texture, and ultimately not a basic mode at all. In other words, the primary goal of the mechanical philosophy was to explain natural phenomena in terms of two basic properties: matter and motion. With the reduction made possible by La Forge's occasionalism, natural phenomena can be explained in terms of just one.

The Cartesian reduction of motion is also important in terms of intelligibility. Although mechanical explanations of the seventeenth century are recognised by historians of science today, primarily for their innovative use of experimentation and the quantification of empirical observation, they were originally championed mostly for their superior intelligibility. Intelligibility was perhaps the most highly prized virtue of the *new* philosophy. An example of Descartes' to support this claim is of special interest given the issue at hand. Consider Descartes' rejection of the Aristotelian conception of motion. Although he makes the same point on different occasions, our favourite comes from the seventh chapter of *The World*:

[Aristotelians] admit themselves that the nature of their motion is very little understood. To render it in some way intelligible they have not yet been able to explain it more clearly than in these terms: *Motus est actus entis in potentia, prout in potentia est*. For me these words are so obscure that I am compelled to leave them in Latin because I cannot interpret them. (And in fact the sentence 'Motion is the actuality of a potential being in so far as it is potential' is no clearer for being translated.) . . . For my part, I am not acquainted with any motion except that which is easier to conceive than the lines of the geometers—the motion which makes bodies pass from one place to another and successively occupy all the spaces which exist in between (CSM I, 93–94/AT XI 39).

Descartes rejects the Aristotelian conception of motion in favour of a definition which he claims is more intelligible, and he seems to be correct. This reflects one of the explanatory goals of Cartesian science and the broader mechanical philosophy, which sought to peel back the layers of nature with deeper and deeper mechanical explanations. For instance, the functions of the hands of a clock can be explained by appealing to the arrangement of the gears, while the gears themselves might be explained in terms of the interaction of microscopic corpuscles. Descartes' explanation of motion is more intelligible than the one he rejects, but it is still problematic. Defining motion as a *passing* from one location to another begs the question and ultimately accepts motion as a primitive. While it is true that all explanations must ultimately end somewhere, La Forge realised that as a Cartesian he was already committed to solid Cartesian principles that could push the explanation of motion even deeper. The passing from one location to another involves only God's continued creation of the moving object at different locations over time.



While La Forge often speaks as if motion is a basic mode, it is clear that he recognised the reduction of modes that this explanation allowed. This is best seen when he writes, ‘For motion, considered in the body which is moved, is only the transfer of a body from the vicinity of those which are in immediate contact with it and which are regarded as being at rest, to the vicinity of other bodies’ (FHM 145/FOP 238).

This reduction of motion also seems to have had some success, being adopted by other important occasionalists. Most significantly, Malebranche, the most comprehensive and systematic of the Cartesian occasionalists, offers an argument very similar to La Forge’s that seems to acknowledge the reduction. Malebranche writes:

Creation does not pass: the conservation of creatures is on the part of God simply a continued creation, simply the same volition which subsists and operates unceasingly. Now, God cannot will that this chair exist and, by this volition, create or conserve it without His placing it here or there or elsewhere. Hence, it is a contradiction that one body be able to move another. I say further: It is a contradiction that you should move your chair. Even that is not enough. It is a contradiction that all the Angels and Demons joined together should be able to move a wisp of straw (MPS 230/OC 12, 161).

Some scholars claim that Malebranche never advocated the cinematic view, but we feel that the similarity of the preceding passage to the one by La Forge quoted earlier, especially given the emphasis on location, indicates that Malebranche was likely influenced by La Forge’s view. However, at least one other Cartesian explicitly accepted this account of motion and clearly recognised the reduction. In the *Physica Electiva* (1697), Johann Christopher Sturm claims that motion ‘is only the successive existence of the moving thing in different places’.<sup>14</sup> Although known only as an obscure Cartesian today, Leibniz found him so influential that he composed one of his greatest works, *On Nature Itself*, as a reply to his views.

The Cartesian reduction of motion promises to offer a high degree of both intelligibility and parsimony to the Cartesian project. But what makes this reduction possible is the commitment to certain solid Cartesian principles, which, considered collectively, entail occasionalism. The next part of the chapter explores occasionalism in more detail and focuses on the Cartesian idea that the modes of a material substance cannot be communicated to any other material substance.<sup>15</sup>

Once again, by the expression *solid Cartesian principle*, we mean any rule or thesis that has a conspicuous association with the Cartesians in general, especially a thesis that has been traditionally regarded as a principle that separates the Cartesians from the Scholastics in a sharp and reactionary way. Moreover, it is a principle that a philosopher who is typically labelled ‘Cartesian’ would be quite reluctant to abandon. There are several



such principles underlying La Forge's account of motion. But now, a point of anticipated contention must be addressed.

As is well known, a good number of Descartes' followers were occasionalists, especially during the second half of the seventeenth century. Moreover, these particular Cartesians did not regard occasionalism as a position that was merely incidental to or logically independent of Descartes' own philosophy. Rather, practically all of these later seventeenth-century Cartesian occasionalists regarded occasionalism as a quite natural, logical development of Descartes' thinking. However, if anyone was a proponent *par excellence* of solid Cartesian principles, it was surely Descartes himself. But, goes the objection, most contemporary scholars of Descartes do not regard him as an occasionalist. Also, if Descartes was an occasionalist, he was certainly not very open about it. This is certainly a point of controversy worth noting.<sup>16</sup>

We feel, on the other hand, that Descartes may very well have been an occasionalist, at least concerning body-body interaction. To defend this point as briefly as we can in this chapter, it will be instructive to look at a discussion of Malebranche by Steven Nadler, and to consider a particular argument that Malebranche attributes to Descartes. The task will be to determine whether Descartes actually espoused this argument, either explicitly or by tacit implication.

In his discussion of one of Malebranche's reasons for denying that bodies are genuinely causally efficacious, Nadler points out that Malebranche held the following three views: (1) if a bodily substance has motive force, then that force would have to be a modification of that particular body; (2) for one body to cause motion in another body, the former must be able to communicate its moving force to the latter; and (3) modes cannot be communicated from one finite substance to another.<sup>17</sup> Obviously the upshot here is that on Malebranche's view one body cannot be the cause of another body's motion. As Nadler points out, the only other option that might be consistent with 'real' bodily causation is the view that bodies cause motion in other bodies by creating, *ex nihilo*, power in those other bodies. However, as explained by Nadler, this opinion is inconsistent with the view that extension is the essence of bodily substances and that extension is essentially passive.<sup>18</sup> As Malebranche puts it, in the Seventh Dialogue in the *Dialogues on Metaphysics*:

Consult the idea of extension and judge by that idea, which represents bodies if anything does, whether they can have some property other than the passive faculty of receiving various shapes and various motions. Is it not evident to the last degree that properties of extension can consist only in relations of distance? (MPS 224/OC 12, 150)

Furthermore, this latter suggestion does not square with the opinion that creation *ex nihilo* does not fall within the capacity of finite substances.



As is well known, Malebranche imputes theses (1), (2), and (3) to Descartes. So part of our task is to muster up textual support that Malebranche was correct to attribute those views to Descartes. Once again, thesis (1) is the claim that if a body has the power to move another body, then that power must be a modification of the former body, and so, by implication, that power must be a mode of extension. By our reckoning, Descartes' commitment to thesis (1) practically goes without saying. For instance, in a letter to Henry More, dated August 1649, Descartes explicitly asserts that in a created substance, the power to cause motion is a mode.<sup>19</sup> Also, in part one of the *Principles*, Descartes makes it clear that any characteristic that can be attributed to a body is a mode of extension.<sup>20</sup>

Therefore, we shall turn our attention to thesis (2), which is the claim that if one body can 'truly' cause motion in another body, then the former body must bear a relation to the latter that involves a genuine transfer of power, a conveyance of an efficacious quality from the former to the latter. It is tempting to presume, based on general reading, that Malebranche was correct to attribute thesis (2) to Descartes. We shall call thesis (2) the *Bodily-Power-Transfer-Thesis* (BPTT). However, some textual support from Descartes himself will be instructive, for in that way we can really justify the view that Descartes endorsed BPTT, rather than merely assuming that he did so. Also, it is important to acknowledge that the Bodily-Power-Transfer-Thesis and thesis (1) are only conditional statements, and, as such, they are impartial to the question of whether bodies are really causally efficacious or not.

But for starters, why is the concept of a 'power transfer' so important to both Malebranche and Descartes? The beginning of a solid answer is Nadler's remark that the logic of causality involves a relationship of necessity between things or events, whereas the metaphysics of causality involves 'power'. As Nadler explains it, these two aspects of causality are closely related by the fact that the necessity of a casual connection is supposed to be determined by a real power in a causal agent. Presumably, both Malebranche and Descartes held this view, and so did Descartes' scholastic contemporaries. To state this view schematically, if a thing, A, is the genuine cause of something, B, then B follows necessarily from A by virtue of the fact that A has a nature that renders it with the power sufficient to bring about B, so long as God permits it. As Nadler puts it, if this view is true, then causal necessity has, as its source, a genuine metaphysical foundation.<sup>21</sup>

We think that there are a number of passages that indicate that Descartes regarded a genuine causal relation as one that involves a transfer of power from one thing to another.

One passage in particular that we find especially suggestive is from the *Third Meditation*, where in an argument for the existence of God, Descartes says that 'it is manifest by the natural light that there must be at least as much <reality> in the efficient and total cause as in the effect of that



cause. For *where, I ask, could the effect get its reality from, if not from the cause?*' (CSM II, 28/AT VII, 40, our italics). The pertinent idea here is Descartes' view that an effect contains something that belonged to or belongs to the efficient cause of that effect, and, furthermore, the formal reality of the effect is numerically identical to some of the metaphysical content that is or was formally contained in the cause.

Our interpretation of the preceding passage is espoused also by Daisie Radner, who argues that Descartes' view on the general nature of efficient causation can be completely characterised by the following three principles:<sup>22</sup>

- (1) *The Communication Principle*: The efficient cause communicates some of *its* reality to the effect.
- (2) *The Pre-Existence Principle*: The reality of the effect is already contained in the cause.
- (3) *The 'at least as much principle'*: The cause must contain at least as much reality or perfection as that which is contained in the effect.

According to Radner, (1) logically implies (2), and (2) implies (3). Also, we would like to mention that Radner's interpretation has been thoroughly criticised by a number of scholars, including Louis Loeb and Tad Schmaltz.<sup>23</sup> A detailed discussion of their objections would require an extended exposition in its own right, one that would go beyond what is practicable in the present chapter. However, we would like to briefly address a particular objection raised by Loeb that happens to be very pertinent to our position.

In support of the communication principle, Radner cites the following question asked by Descartes: 'And how could the cause give it [reality] to the effect unless it possessed it?' (CSM II, 28/AT VII, 40).<sup>24</sup> Loeb points out that Radner's argument requires the presumption that (1) the verb 'give' in this passage should be translated as 'communicate' and (2) in this context the verb 'communicate' should be understood in the sense of 'transfer'. Loeb goes on to argue that Radner has not provided sufficient evidence to support the translation of the Latin *dare* here as 'to transfer'. Loeb correctly points out that the present active infinitive *dare*, commonly translated as 'to give', is not one of the most obvious Latin terms for 'to transfer'. He concludes by saying that the infinitives 'to give' and 'to communicate' can be used to mean *to cause to have*, without specifically meaning 'to cause to have *by means of transference*'.<sup>25</sup>

We believe that Loeb's argument is basically correct, and that the passage Radner cites from the Third Meditation does not conclusively support causal principle (1). However, there is a passage from the Third Meditation that we believe does support (1). Right after Descartes asserts the claims that Radner believes count as evidence for Descartes' commitment to (1), (2), and (3), Descartes states that the cause of a particular idea must have at least as much formal reality as there is objective reality in that idea.<sup>26</sup> What we find interesting is Descartes' subsequent clarification that the efficient



cause of the existence of a certain idea ‘does not transfer any of its actual or formal reality’ to the idea itself.<sup>27</sup> For quick reference, let us call the former view ‘Principle D’. One thing deserving of attention here is Descartes’ use of the verb ‘transfer’ in this passage, which is translated from the verb ‘*transfundat*’, the third-person present subjunctive form of the infinitive ‘*transfundere*’. The two most common meanings of ‘*transfundere*’ are ‘to transfer’ and ‘to pour *from* one vessel *into* another’. By our reckoning it seems very unlikely that Descartes would have thought that it was worthwhile to state Principle D if he had not thought that it was true without exception that no individual substance, *A*, transfers any of its formal reality to another thing, *B*, unless *A* stands to *B* as efficient cause to effect.

Concerning causal relations among bodies via collisions, in Part Two of the Principles, article 40, where Descartes articulates an argument for his third law of motion, he states that if a body collides with a smaller body, the quantity of motion that the larger body loses is ‘imparted’ to the smaller body (See CSM I, 242/AT VIII A, 65). The idea to notice here is Descartes’ explanatory idea of one body transferring a power or motion-causing quality to another body. Even if Descartes didn’t intend for the notion of a ‘power transfer’ to be taken literally when it comes to explaining the communication of motion among bodies, he still must have thought that it was at least a convenient and useful idea to employ in the context of an ordinary physical explanation.<sup>28</sup>

Moreover, in a letter to Henry More, dated August 1649, Descartes says that he agrees with More’s contention that if matter is left entirely undisturbed and does not receive an impetus from anything, then it will remain completely undisturbed. He continues by specifically saying that ‘it [matter] receives an impulse from God, who preserves the same amount of motion or transfer in it as he placed in it at the beginning’.<sup>29</sup> Here we see the idea that matter must receive into it some kind of power or force in order for particles to be set into motion. At any rate, we conclude that there is enough textual support to be confident that Descartes was not inclined to question the idea that a causal relation is one that involves the transfer of some kind of power from one thing to another.

Now we come to consideration of thesis (3), the view that Cartesian metaphysics precludes modes from being conveyed from one finite substance to another. We wish not to defend this claim in all of its generality, but to defend a restricted version of it, namely, what thesis (3) entails when relativised to bodies. But where is the relevant textual evidence to be found that Descartes accepted a body-restricted version of (3)? Before answering this question, we would like to turn briefly, once again, to Malebranche, for it is he who would have us believe that thesis (3) is an obvious product of Descartes’ thinking.<sup>30</sup> In his response to a criticism made by Antoine Arnauld to a clarification of the *Treatise on Nature and Grace*, Malebranche makes the following remark: ‘If moving force belonged to bodies, it would be a mode of their substance, and it is a contradiction that modes go from substance to substance.’<sup>31</sup> Let



us call the view that modes cannot be communicated across substances the *No-Transfer-Thesis*. Malebranche and others were strongly convinced that Descartes was committed to the No-Transfer-Thesis, but they rarely cite any explicit textual evidence for this contention.

Admittedly, candid textual evidence indicating a commitment to the No-Transfer-Thesis is not easy to find in the writings of Descartes. But we think that there is at least one set of remarks from Descartes which fairly clearly and unequivocally indicate his commitment to a body-restricted version of the No-Transfer-Thesis. The passages we have in mind can be found in a letter to Henry More of August 1649. In that particular letter, Descartes responds to More's complaint that Descartes' principles of physics entail that numerically one and the same motion or force can successively occupy a train of smaller and smaller bodies. Now, in this letter Descartes explicitly states that motion is a mode and not a subsistent entity that comes to occupy different locations while remaining numerically one and the same entity. Furthermore, Descartes says that More was mistaken when he attributed to him the view that numerically one and the same mode can successively occupy a train of smaller and smaller bodies. We shall explain this point in what is to follow.

Concerning the view that motion cannot be communicated from one body to another, Descartes tells More that he, Henry More, observes correctly that 'motion, being a mode of body, cannot pass from one body to another' (CSM III, 382/AT V, 404). So it seems that the explanation that Descartes presents to More for denying the transfer of motion is the same reason that Malebranche espouses and imputes to Descartes: motion is a modification of bodies. Descartes' argument here is simple and enthymematic, with the tacit premise clearly being the claim that modes of bodily substances are not capable of being transferred from one body to another.

In the same letter, Descartes goes on to outline what he takes to be the correct way to explain what occurs when one body appears to alter the location or acceleration of another body. It should be noted that Descartes' explanation seems to be designed to explain all motion-related modes in terms of nothing more than both contact between bodies and a change in distance relations between a certain body and others presumed to be at rest.<sup>32</sup> For our purposes, what should be acknowledged here is the fact that this particular account of motion-related phenomena does not make use of the idea that bodies actually impart or receive forces or other modes when they undergo the process of collision. This was certainly not a new idea for Descartes, for in the *Principles of Philosophy*, published five years before the letter to More of August 1649, Descartes offers essentially the same account of motion.<sup>33</sup> Also, Descartes' explanations in both the *Principles* and this letter to More clearly indicate that he was not primarily concerned with offering an account of locomotion or continuous change in location. Rather, his intention was to reduce all force-related modes, such as power and momentum, to both contiguity and change in distance relations between neighbouring bodies.



Eventually Descartes assures More that ‘there is no need for you to worry about the transmigration of rest from one object to another, since not even motion, considered as a mode which is the contrary to rest, transmigrates in that fashion’ (CSM III 382/AT V, 405). Since this argument was given toward the end of Descartes’ life, the opinion that he is defending here was probably well considered, and so should be taken seriously. Though one feels that more textual support is needed to make our case compelling, we believe that we have a decisively good start at showing that Descartes was a proponent of the No-Transfer-Thesis.

At this point, we don’t exactly have what Malebranche may have needed to support his contention that Descartes’ philosophy justifies general occasionalism. But, we do have just enough to make plausible our claim that Descartes was an occasionalist concerning body-body causation, and we can use nothing more than theses (1) and (2), and the appropriately restricted version of (3). Articulated into a choppy paraphrase, the argument runs like this:

- (1) All motion-related qualities of bodies are modes. (solid Cartesian principle)
- (2) Every instance of ‘genuine’ causation involves the transfer of some kind of power from one entity to another. (solid Cartesian principle)<sup>34</sup>
- (3) No mode of extension can be transferred from one body to another. (No-Transfer-Thesis as restricted to bodies)
- (4) So, no motion-related quality of bodies can be transferred from one body to another. (From 1 & 3)
- (5) Hence, no instance of genuine causation can involve the transfer of a motion-related quality from one body to another. (From 2 & 4)
- (6) But, a body can causally affect another body only if it can transfer some of its motion-related qualities to the other body. (From 2)
- (7) Therefore, when a body, A, collides with another body, B, A is not the real cause of any change that B undergoes as a result of their collision. (From 5 & 6)

But what now are Descartes’ options? One possibility is that there are cases where finite minds are causally responsible for pushing bodies around to some extent, however small the distance. This much he explicitly says in the same letter to More, which was quoted earlier. But what about all those circumstances where no finite mind is involved in any changes that occur to a body, B, whenever it is impacted by another body, A? The only correct answer here for Descartes, if our hypothesis is correct, is that God is the only causal agent that can make A and B undergo the appropriate changes as a result of their impact. As Descartes himself says in the same letter to More:

The power causing motion may be the power of God himself preserving the same amount of transfer in matter as he put in it in the first moment of creation; or it may be the power of a created substance, like our



mind, or of any other such thing to which he gave the power to move a body. In a created substance this power is a mode, but it is not a mode in God. (CSM III, 381/AT V, 403–4)

Incidentally, one curiosity of the previous passage, and one that is completely consistent with our view, is that Descartes does not explicitly mention bodies as the other candidates that are capable of moving bodies. At any rate, if our interpretation of Descartes is correct, even though finite minds can agitate bodies to some degree, for the rest of the way God is the one who must carry the ball. Another interesting implication of our interpretation is that in Descartes' universe, immaterial entities are the only causally efficacious substances.

It is not practicable within the scope of this chapter to discuss every recently defended view that disagrees with our position on a fundamental level.<sup>35</sup> However, one opposing view that we would like to discuss briefly is the position defended by Helen Hattab. Hattab argues that most scholars read Descartes as being committed to a certain model of causation that was widely held among his Scholastic-Aristotelian predecessors, a model according to which the relation of efficient causation is grounded in certain properties that inhere in either the individual substance that serves as the cause of a particular effect or the thing that happens to be acted upon.

Hattab holds that this interpretation of Descartes is wrong, and argues that, by the time Descartes had completed *Le Monde*, he completely rejected the Scholastic-Aristotelian model of efficient causation. With this model out of the picture, Hattab imputes to Descartes the view that efficient causation is constituted by the following two metaphysical conditions: (1) the immutable decrees of God, which provide the very power by which bodies are conserved in certain ways and conserved in certain relations to one another, and (2) the various features that are fundamental to the nature of bodies and the geometric relations they bear to one another. More specifically, Hattab argues that Descartes' decided view is that the laws of nature themselves are efficacious and that they function as 'the secondary and particular causes of motion'.<sup>36</sup> As Hattab puts it:

The causality of the laws of nature consists not in any intrinsic powers or actions inhering in them as modes. Rather what forms the basis of the laws of nature is another set of relations, the most obvious one being the relation between God's immutable action and the states of matter. But as we have seen, the states of matter are not fully determined by the simplicity and immutability of God's action, rather, God's action has to be accommodated to certain principles basic to the nature of matter.<sup>37</sup>

We find Hattab's reading of Descartes interesting and well defended, and we recognise that in Part Two of the *Principles*, Descartes does say that the laws of nature are 'the secondary and particular causes of the various



motions we see in particular bodies'.<sup>38</sup> However, we have two reservations about her interpretation.

One apparent problem with Hattab's interpretation is due to the fact that, as we have already seen, in a letter as late as 1649, Descartes unequivocally asserts that in a finite substance, the power to cause motion is a mode. Our main concern, however, is that it is not entirely clear to us how the position Hattab imputes to Descartes is to be distinguished from Malebranche's view that bodies are completely devoid of motor power and that God is the only entity capable of moving bodies. For instance, in the *Dialogues on Metaphysics*, the newly convinced Aristes concurs with Theodore that 'God communicates His power to us only through the establishment of certain general laws, the efficacy of which we determine by our various modalities.'<sup>39</sup> The import of this particular passage from Malebranche strikes us as being very close to the sort of position that Hattab attributes to Descartes. It seems to us that what Hattab refers to as the states of matter are really nothing more than Malebranche's occasional causes. The language is different, but the implications are the same. Since it is certainly not Hattab's intention to hold that body-body occasionalism is an implication of Descartes' metaphysics, we believe that there is a strong need here to more clearly articulate how the bodies of Descartes' metaphysics are to be separated from those of Malebranche's philosophy.

We hope to have shown in our chapter why the seventeenth-century Cartesian occasionalists believed that their brand of physical mechanics satisfied the explanatory demands of parsimony and intelligibility much better than any other approach. In particular, Cartesian mechanics was thought to achieve such explanatory excellence by accounting for motion in a way that reduces the many modes of extension to the smallest number possible. The solid Cartesian principles required for this reduction entail occasionalism. Being solid Cartesian principles, it seems likely that Descartes recognised this simple entailment and consequently accepted occasionalism, at least regarding body-body causation.

As our conclusion would have it, the Cartesians who were proponents of occasionalism are in a class that includes Descartes himself. But it certainly doesn't include every faithful Cartesian. Pierre Régis and Jacques Rohault were, for all practical purposes, defenders of 'solid Cartesian principles', yet they were not proponents of occasionalism. Then, there was Antoine Arnauld, who affirmed the existence of both body-body and mind-body causation, but denied that bodies were capable of affecting immaterial substances. This, however, is a topic for another occasion.

## NOTES

1. Abbreviations Used: AT: R. Descartes, *Oeuvres de Descartes*, edited by C. Adam and P. Tannery (Paris, 1897; reprint, Paris: J. Vrin, 1964–1975). Cited by volume and page. CSM: R. Descartes, *The Philosophical Writings of*



- Descartes*, translated by John Cottingham, Robert Stoothoof, and Dugald Murdoch, and Anthony Kenney, 3 volumes (Cambridge: Cambridge University Press, 1985). Cited by volume and page. FHM: L. La Forge, *Treatise on the Human Mind*, translated by Desmond Clarke (Dordrecht: Kluwer Academic Publishers, 1997). FOP: L. La Forge, *Oeuvres Philosophiques*, edited by P. Clair (Paris: Presses Universitaires de France, 1974). MPS: Malebranche, N. *Malebranche: Philosophical Selections*, edited by Steven Nadler (Indianapolis/Cambridge: Hackett Publishing Co., 1992). OC: N. Malebranche, *Oeuvres Complètes de Malebranche*, edited by André Robinet, 20 volumes (Paris: J. Vrin, 1958–67). Cited by volume and page.
2. See Daniel Garber, “How God Causes Motion: Descartes, Divine Sustenance and Occasionalism”, in *Descartes Embodied: Reading Cartesian Philosophy through Cartesian Science*, edited by Daniel Garber (Cambridge: Cambridge University Press, 2001), 189–202. See also Helen Hattab, “Concurrence or Divergence? Reconciling Descartes’ Physics with His Metaphysics”, *Journal of the History of Philosophy*, 45 (2007) 1: 49–78. See also Tad Schmaltz, *Descartes on Causation* (Oxford: Oxford University Press, 2008).
  3. Descartes’ frequent reference to the natural light of reason can be used as a guideline for identifying solid Cartesian principles. If Descartes thinks a view is justified by the natural light of reason, it’s usually safe to assume it’s a view he would be reluctant to reject.
  4. This succinct formulation is due to Robert Sleight. See Robert C. Sleight Jr., “Leibniz on Malebranche on Causality”, in *Central Themes in Early Modern Philosophy: Essays Presented to Jonathan Bennett* (Cambridge: Cambridge University Press, 1990), 161–93.
  5. See S. Nadler, “The Occasionalism of Louis de la Forge”, in *Causation in Early Modern Philosophy: Cartesianism, Occasionalism, and Pre-established Harmony*, edited by S. Nadler (University Park, Penn.: Penn State University Press, 1993), 57–73.
  6. A similar claim is made in the *Principles of Philosophy*, Part One, Article 21 (See CSM I, 200/AT VIII A, 13).
  7. See Schmaltz, *Descartes on Causation*, 102–3.
  8. See Schmaltz, *Descartes on Causation*, 106–16.
  9. Walter Ott writes, “To Talk of a body continuing along in a given direction and speed ‘quantum in se est’ is only a way of saying that the body will do so, in virtue of God’s nature and will”, *Causation and Laws of Nature in Early Modern Philosophy* (Oxford: Oxford University Press, 2009), 57–58.
  10. This interpretation seems compatible, for instance, with the Part One, Section 26 of the *Principles of Philosophy* (CSM I, 201/AT VIII A, 14).
  11. For an account of Boyle’s mechanical affections see William R. Eaton, *Boyle on Fire: The Mechanical Revolution in Scientific Explanation* (London: Continuum, 2005), 134–146.
  12. Law #1: What is once in motion always continues to move. Law #2: All motion is rectilinear. Law #3: If a body collides with another body that is stronger than itself, it loses none of its motion; but if it collides with a weaker body, it loses a quantity of motion equal to that which it imparts to the other body.
  13. Robert Boyle, *Works of Robert Boyle*, Vol. 8, edited by Michael Hunter and Edward Davis (London: Pickering and Chatto, 2000), 106.
  14. See Leibniz, *Philosophical Essays*, translated by Roger Ariew and Daniel Garber (Indianapolis/Cambridge: Hackett, 1989), 163. Having only recently discovered the significance of Sturm’s occasionalism to this project, a more detailed analysis of his philosophy is in development.
  15. In “Cordemoy and Occasionalism”, *Journal of the History of Philosophy*, 43 (2005) 1: 37–54, at pp. 37–38, Steven Nadler mentions the fact that there are



two general Cartesian principles that stand as obstacles to justifying the view that bodies possess causal powers: (1) the nature of matter is nothing more than extension and (2) the modes of one substance cannot be communicated to another substance.

16. For a survey of the recent scholarship on this issue, see Helen Hattab, "Concurrence or Divergence? Reconciling Descartes' Physics with his Metaphysics". See also Tad Schmaltz, *Descartes on Causation*.
17. See Steven Nadler, "Malebranche on Causation", in *The Cambridge Companion to Malebranche* (Cambridge: Cambridge University Press, 2000), 120.
18. Nadler, "Malebranche on Causation", 120.
19. See CMS III, 381/AT V, 403–4.
20. See CMS I, 210/AT VIIIA, 210.
21. For a great discussion of this matter, see Steven Nadler, "Malebranche on Causation", 115. Concerning the significance of the concept of necessity in what may have been the standard seventeenth-century understanding of causality, we concur with Robert Sleigh ("Leibniz on Malebranche on Causality", 171), according to whom the rationalists in general were committed to the following generic conception of causation:  $x$  is a real cause of  $E$  only if there is some state  $C$  of  $x$ , such that, necessarily, if  $C$  obtains then so does  $E$ . Not surprisingly, perhaps, Malebranche thought that metaphysical necessity is the only type of necessity appropriate for qualifying a relation as causal in nature.
22. See Daisy Radner, "Is There a Problem of Cartesian Interaction?", *Journal of the History of Philosophy*, 23 (1985) 1: 35–49.
23. See Louis Loeb, "Replies to Daisy Radner's 'Is There a Problem of Cartesian Interaction?'", *Journal of the History of Philosophy*, 23 (1985) 2: 227–31. Also, a very recent and informative discussion of this very issue and the exchange between Radner and Loeb can be found in the work of Schmaltz's, *Descartes on Causation*, 49–86. Also, see Tad Schmaltz, "Deflating Descartes' Causal Axiom", *Oxford Studies in Early Modern Philosophy* (Oxford: Clarendon Press, 2006) 3: 1–31.
24. See Radner, "Is There a Problem of Cartesian Interaction?", 41.
25. Loeb, "Replies to Daisy Radner's 'Is There a Problem of Cartesian Interaction?'", 227.
26. See CSM II, 28–29/AT VII, 41. Radner regards this passage as evidence for Descartes' commitment to the "at least as much" causal principle.
27. CSM II, 28–29/AT VII, 41.
28. Descartes discusses the distinction between the ordinary sense of the term 'motion' and the technically true sense of 'motion' in articles 24 and 25 of Part Two of the *Principles*. See CSM I, 233/AT VIIIA, 53–54.
29. CSM III, 381/AT V, 404: 'Illa . . . impellitur à Deo, tantumdem motûs siue translationis in eâ conseruante, quantum ab initio posuit.'
30. Malebranche was certainly not alone with this contention, for La Forge also believed that Descartes was committed to the No-Transfer-Thesis. As we have very recently discovered, La Forge's evidence for Descartes' commitment to the No-Transfer-Thesis is the same as our evidence, namely, Descartes' letter to Henry More, of August 1649. See FHM, 149.
31. Réponse à une Dissertation de Mr. Arnauld contre un Eclaircissement du Traité de la Nature et de la Grace, VII, 6, OC 7, 515–16.
32. See CSM III, 382/AT V, 404–5.
33. See CSM I, 233/AT VIIIA, 53–54.
34. The reader should note that the second assumption of the preceding argument entails the Bodily-Power-Transfer-Thesis (BPTT), which is discussed on



pages 54–57. BPTT is only conditional in character. For our purposes, we have chosen to use a completely unrestricted thesis concerning the transfer of power in a causal relation. Concerning Descartes' endorsement of this unrestricted causal principle, see pp. 54–57 of the present article. Some scholars, however, may just take it as a platitude that Descartes was committed to the second assumption. At any rate, we feel confident enough about Descartes' commitment to the second assumption to label it as a solid Cartesian principle.

35. For instance, Schmaltz argues that Descartes' official position is that bodies possess a causal efficacy distinct from divine volitions and that a body's power to cause motion is grounded in the nature of corporeal substance. See Schmaltz's, *Descartes on Causation*, and "Deflating Descartes' Causal Axiom". On the other hand, Peter Machamer and J. E. McGuire have recently defended a view that is compatible with our position. See P. Machamer and J. E. McGuire, *Descartes' Changing Mind* (Oxford and Princeton, NJ: Princeton University Press, 2009).
36. Helen Hattab, "Conflicting Causalities: The Jesuits, Their Opponents, and Descartes on the Causality of the Efficient Cause", in *Oxford Studies in Early Modern Philosophy*, edited by D. Garber and S. Nadler (Oxford: Clarendon Press, 2003) Vol. 1: 19.
37. Hattab, "Conflicting Causalities: The Jesuits, Their Opponents, and Descartes on the Causality of the Efficient Cause", 21.
38. CSM I, 240/AT VIIIA, 62.
39. N. Malebranche, in *Malebranche: Philosophical Selections*, edited by S. Nadler (Indianapolis/Cambridge: Hackett, 1992), 235.



# 4 Spinoza's Conatus as an Essence-Preserving, Attribute-Neutral Immanent Cause

## Toward a New Interpretation of Attributes and Modes

*Eric Schliesser*

To José Benardete and Kris McDaniel<sup>1</sup>

### I. THESIS & INTRODUCTION

In this chapter I characterise exactly what kind of cause Spinoza's *conatus* is. I do this in two steps. First, I ask and answer the following question: 'What philosophical problem could Spinoza's *conatus* concept be seen as solving?' This illuminates the nature of the concept and, given its central importance to Spinoza, thus, a core aspect of Spinoza's system. By reflecting on some of Spinoza's youthful observations (on Descartes), I argue that Spinoza's monism can be fruitfully understood as responding to a number of interconnected problems in Descartes' treatment of divine concurrence and conservation (Sections II–IV).<sup>2</sup> To be clear, Spinoza adapts a Cartesian framework for his own very un-Cartesian ends.<sup>3</sup> I assume without further argument that Descartes' metaphysics is designed to make natural philosophy possible,<sup>4</sup> while Spinoza's metaphysics intends to make enduring happiness possible.

I claim that for Spinoza *conatus* characterises an attribute-neutral, essence-preserving immanent cause.<sup>5</sup> In order to understand what this could mean, I offer, second, a commentary on the details of Spinoza's demonstration of EIIIp6.<sup>6</sup> By reinterpreting some central Spinozistic concepts, 'attributes', 'expression', and 'modes', I point to a radically different approach to Spinoza's metaphysics than common in recent scholarship (Section V). In particular, I reject the tendency to read the *conatus* doctrine as some articulation of a variant of the mechanical philosophy. To prefigure one of my conclusions: the essences that are being preserved are not located in space and time.

Thus far in the scholarly literature, understanding Spinoza's *conatus* concept has been tackled from three main directions (with hybrids among



them): (1) the proof of EIIp6 has been closely scrutinised and been found wanting by influential commentators,<sup>7</sup> although there have been defenders;<sup>8</sup> (2) many have looked at Ancient,<sup>9</sup> Hellenistic,<sup>10</sup> Medieval,<sup>11</sup> and early modern<sup>12</sup> sources for and predecessors to a *conatus* doctrine; (3) others have explored how Spinoza applies the *conatus* doctrine, especially in his psychological<sup>13</sup> and political doctrines.<sup>14</sup> Thus far, I am unaware of investigations into what perceived defects the *conatus* doctrine could be redressing (but Lachterman as well as Garber offer useful studies).<sup>15</sup> Most of my treatment focuses on (difficult aspects of) the text of the *Ethics*, although I also call attention to an illuminating passage from Bacon's *New Organon*. A pay-off of my approach of looking at Spinoza's commentary on *Descartes' Principles of Philosophy* and his *Metaphysical Cogitations* is that it opens the way to a developmental understanding of Spinoza's thought.

## IIA. Spinoza's *Conatus* Doctrine

Early in the third part of the *Ethics*, Spinoza states the *conatus* doctrine and demonstrates it as follows:

EIIp6: Each thing, naturally [*quantum in se est*]<sup>16</sup> strives [*conare*] to persevere in its being.

Dem: [A] For singular things are modes by which God's attributes are expressed in a certain and determinate way (by 1P25C), that is [B] (by IP34) things that express, in a certain and determinate way, God's power, by which God is and acts. [C] And no thing has anything in itself by which it can be destroyed, or which takes its existence away (by P4). [D] On the contrary, it is opposed to everything which can take its existence away (by P5). Therefore, as far as it can, and it lies in itself, it strives to persevere in its being, q.e.d.

My thesis that Spinoza's *conatus* doctrine is meant to describe an attribute-neutral, essence-preserving immanent cause has considerable initial plausibility. EIIp6 states that each entity strives to persevere in its being. One could simply assert that *conatus* just is a way to speak of an entity's power to persevere. This is less strange than it sounds, given that Spinoza rejects final causes (see, for example, the influential appendix to *Ethics* 1);<sup>17</sup> it is a 'non-goal-oriented' striving. Such 'striving' must be a kind of causation that is not oriented toward a goal, as in final causation, but rather is 'directed' by its source. Besides Spinoza's rejection of final causation, there is very strong evidence for this claim because in the very next proposition, EIIp7 ('the striving by which each thing strives to persevere in its being is nothing but the actual essence of the thing'), *conatus* is identified with the actual essence of a thing.<sup>18</sup> So, the actual essence of a thing is identified with the way a thing perseveres. I explicate this claim in Section V.



Nonetheless, even if we put aside EIIIP7, the demonstration to proposition EIIIP6 suggests that things cannot be so simple, especially because, although Spinoza defines an essence at the start of part II of the *Ethics*,<sup>19</sup> and this definition entails that an essence is a necessary condition for the thing's existence of which it is an essence, it is far from clear what essences are in Spinoza. The demonstration introduces at least four different claims, which in the block quote earlier I have labelled 'A' through 'D' for ease of exposition. I need to articulate how C and D are best interpreted as helping to explain how *conatus* is 'essence-preserving', and explain how A and B show that it is illuminating to speak of *conatus* as an attribute-neutral and immanent cause. Now the most direct interpretive route lies in explaining what Spinoza means by 'attributes', 'modes', and 'expression', but I resist that temptation initially in order to provide some useful background. I pass, first, through some familiar elements of Descartes' metaphysics and then I call attention to some of Spinoza's remarks on these.

## IIB. Detour Through Descartes

My claims about Descartes should be familiar to scholars and teachers of early modern philosophy. For the sake of clarity and brevity, I pretend there are no interpretive debates over these.

### IIBa. Descartes' Definitions of Substance

As is well known, Descartes provides three not entirely compatible definitions of 'substance':

1: Substance . . . [is] a thing which exists in such a way that it needs no other thing in order to exist . . . [namely] God. We perceive that, on the contrary, all others can exist only with the help of God's concurrence [*concursus*] . . . the term 'substance' does not apply . . . univocally . . . each substance has one principal attribute (*Principles*, I.51 & 53).<sup>20</sup>

2: 'Substance' . . . applies to everything in which whatever we perceive immediately resides, as in a subject, or to everything by means of which whatever we perceive exists. . . . The only idea we have of a substance itself, in the strict sense, is that it is the thing in which whatever we perceive (or whatever has objective being in one of our ideas) exists, either formally or eminently . . . a real attribute cannot belong to nothing. . . . The Substance which we understand to be supremely perfect, and in which we conceive absolutely nothing that implies any defect or limitation in that perfection, is called 'God' (*Meditations, Second set of Replies, "Definitions"* V & VIII, AT 161).



3: all substances, or things, which must be created by God in order to exist, are by their nature incorruptible and cannot ever cease to exist unless they are reduced to nothingness by God's denying his concurrence to them (*Meditations*, Synopsis, AT 14)

Descartes admits he uses 'substance' in two ways: first, to pick out a supremely perfect, self-sufficient 'God' in which other things are perceived to exist (either formally or eminently); second, to discuss created substances, which are incorruptible and never cease to exist, in which other things are perceived to exist (either formally or eminently). Such created substances have principal attributes and require God's concurrence for their existence.<sup>21</sup> Descartes identifies two such created incorruptible substances: body (*res extensa*) and mind (*res cogitans*), their principal attributes being extension and thought.

Now even though the only idea we have of substance in the 'strict sense' is as the 'thing in which whatever we perceive (or whatever has objective being in one of our ideas) exists', Descartes assigns the supremely perfect substance an important metaphysical role: it concurs in the existence of created substances.

Here I do not explore to what degree Descartes was a full-blooded occasionalist and if he denied the efficacy of second causes (i.e., the laws of nature). I note two features of the Cartesian argument for divine concurrence. First, this argument relies on what we might call the joint operation of a 'separability' thesis and a thesis about God's supreme power, that is, if two things can be clearly and distinctly conceived apart then they can be made to exist apart by God. Descartes mentions these in a well-known passage from the *Meditations*: 'I know that everything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it. Hence the fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct, since they are capable of being separated . . . ' (*Meditation VI*, AT VII, 78). These two theses are presupposed by Descartes in a very intricate argument about the lack of mutual dependence of each part of time.<sup>22</sup> I quote the full passage before commenting on it:

The parts of time are not mutually dependent, and never coexist . . . from the fact that we now exist, it does not necessarily follow that we shall exist a moment from now, unless some cause (that is the same which first produced us) shall, as it were, continually reproduce us, that is, conserve us. For we easily understand that *there is in us no power by which we may conserve ourselves*. And that He in whom there is so much power that He can conserve us separately from Himself, must also conserve himself all the more, or rather, must require no conservations by anyone, and finally, must be God (*Principles*, I.21; see also "Replies to Objections to Meditation III" [emphasis added—ES]).



Descartes' God is the first cause of things and self-conserving. The same cannot be said about created entities because they exist in time. But because moments of time are separable, and given the 'separability' thesis, they require some connecting principle or cause to ensure that created entities exist from one moment to the next.<sup>23</sup> Descartes has God do this work. So, Descartes' God is not only the first cause of things and the concurring cause of created eternal substances; he is also the conserving cause of individual entities that we perceive to exist (in created eternal substances).<sup>24</sup> For my current purposes it leads too far astray to investigate further the relationship between God's concurrence and God's conservation in Descartes. Nonetheless, it is clear that, in Descartes, created entities require 'outside' help to maintain their identity over time. Their dependence on God helps explain their temporal unity.

At this stage one might expect that I now quickly and temptingly claim that Spinoza's *conatus* 'naturalizes' the role of God in preserving entities over time. While this is not completely mistaken, it glosses over some important subtleties.

Moreover, there is a natural objection to the thought that God plays a crucial role in entity preservation that will occur to most readers of Descartes. He has some famous principles, the laws of nature, which may seem to do the (naturalistic) conceptual work claimed to be missing earlier. Descartes' first law of nature is famously formulated as follows: 'Each thing [*res*] naturally [*quantum in se est*] always remains in the same state [*statu*] . . . Hence what is in motion always continues to move' (*Principles*, II.37). The second law ('all motion is in itself rectilinear') is explained as 'Every piece of matter [*partem materiae*], considered in itself, always tends to continue moving, not in any oblique path but only in a straight line' (*Principles*, II.39). One might argue that these two laws of motion show that Descartes does have 'naturalistic' resources to account for the preservation of things over time. This is especially plausible if one accepts (as I do) that Descartes holds that these laws of nature are (what he labels) 'second causes' that operate within nature. Regardless of how Descartes reconciles God's conserving and concurring role(s) with his embrace of the reality of second causes,<sup>25</sup> it might appear misleading to say that Descartes has no naturalistic resources to account for preservation over time. The first law (while of course derived from metaphysics) does not appeal to God, but simply asserts self-preservation as a fact of nature. Moreover, this objection fits nicely with recent work that has perceptively treated Descartes as a creative modern reinterpreter of Epicurean doctrines (see Wilson). Fair enough. This is why I did not immediately turn to Spinoza.

But note four things. First, the laws *tacitly assume* that entities have an *internal* identity-preserving cause.<sup>26</sup> They do not promote a strategy to discover what can be such a cause. Insofar as Cartesian physics aims to be explanatory, it must presuppose (as a constitutive principle) that it is deviations from states of affairs that promote the search for causes *external* to



the entity (recall the first law).<sup>27</sup> Here Descartes is remarkably silent about the *internal* cause of identity over time.<sup>28</sup> Of course, one might claim that elsewhere Descartes points to God's unchanging nature (and his re-creation of each body at each moment) as a source of natural stability,<sup>29</sup> but (from the vantage point that motivated our turn to the laws) this gives God too large a role and second causes too little. Second, the laws do not spell out if this internal cause of stability should be understood as an (active) resistance to change or some cohesive principle (or both, etc.). Third, it is only about 'simple' and 'undivided' things that we can presuppose such a state-preserving cause. While I am not entirely sure why Descartes is unwilling to extend it to compound, divisible entities, I suspect his 'separability' thesis has something to do with it. For it suggests that if things can be separated we cannot assume that they will stay stable; they require some cause to ensure their connectedness.<sup>30</sup> Finally, Descartes has an unacknowledged slide between speaking of 'things' generally (in law 1) to speaking of 'matter' (law 2). One can excuse the slide within the first law (where the final clause of the law can be taken as a mere illustration). But the second law is only formulated in terms of one created substance, *res extensa*. To put this starkly: Descartes is entirely silent about how to think about 'inertia' in *res cogitans*.<sup>31</sup> So, the first law is either stated too generally or Descartes has left out an important part of the story.

Now with regard to the first three of these points, one might object that the 'force of rest' is the missing natural case of internal stability over time. As Lachterman has pointed out, it serves 'at least three distinct physical roles, namely, to explain (1) the perseverance of a body at rest in its state of rest, (ii) the capacity of a resting being to resist being set in motion by another moving body, and (iii) the cohesion between parts of a solid body.'<sup>32</sup> But while it supplies Descartes with an explanatory principle, it makes matters worse by combining too many functions. That is, a difference without an explanation violates a principle of sufficient reason; a principle that attempts to account for disparate phenomena must be the same principle or it is no explanation at all (we might say it violates a principle of economy).<sup>33</sup>

Here we are not concerned with how Descartes would resolve these issues. Rather, the point is that even if we restrict contextual analysis of possible ways of thinking about *conatus* to a Cartesian background, we need to distinguish between (i) interpreting *conatus* as a kind of God-replacing naturalistic, internal *identity*-preserver of entities over time and (ii) interpreting *conatus* as a kind of external, 'inertial' *state*-preserving principle (see Garber, Vampoulis, and Carrierio, who explicitly compare Spinozistic *conatus* to Cartesian inertia).<sup>34</sup> (i) treats *conatus* as a kind of cohesive principle resisting internal decomposition; (ii) treats *conatus* as resistance to changes caused from without, so that when things deviate from remaining as they are, there must be a cause for this, and that can only be external. I first reject (ii) then the usual way of reading of (i).



### III. YOUNG SPINOZA

In this Section I introduce two texts from Spinoza's *Metaphysical Thoughts*; these were added as a kind of appendix to his commentary in geometric fashion of *Cartesian Principles of Philosophy*.

1: . . . whether His concurrence, by which separate objects are preserved, is the same for all things. Concerning such questions . . . we have no certain knowledge. Yet, in the same way, we very certainly know that this concurrence of God, if it is correlated with his omnipotence, must be unitary, although its effect is manifested in various ways (*Cogitata Metaphysica*, part 2, chapter 7, P7).

2: P1: . . . each single moment God creates things as if anew. From this we have shown that objects have no power of self determination or of operation in themselves. . . . Nevertheless, since there are many who admit and believe in this conserving power of God, but in a different sense from us, we shall recall what has already been proven in order that we may detect this fallacy. We have already clearly shown that present time has no connection with future time (Vid. Ax. 10, Pt. 1.) . . .

P2: When speaking of motion this same question appeared . . . God constantly preserves the same amount of motion in nature. If, therefore, we consider the total amount of matter in motion nothing is added. *It does not seem, however, that the same thing can be said of mental phenomena* . . . finally, since the parts of duration do not have a casual connection, we speak more truly to say then that God continually procreates than to say that he conserves them . . . (*Cogitata Metaphysica*, part 2, chapter 11, emphasis added).

In the first block quote, Spinoza expresses ignorance about the univocity of God's concurrence. A concern he may have in mind is the following: according to Descartes there are two kinds of created eternal substances (*res extensa* and *res cogitans*), which are the bearers of properties (etc.). But in the Cartesian scheme there are no resources to investigate how or if God's concurrence with each substance proceeds in the same fashion and with same principles (etc.). Does God's concurrence operate the same way with mental and extended substances? Descartes is silent about these matters. Even if one is not committed to a full, metaphysically robust version of the principle of sufficient reason,<sup>35</sup> there appears to be a potentially unaccountable difference without a cause at the heart of Descartes' system. Spinoza resolves the problem by letting Divine concurrence be unitary, although (as he says) 'its effect is manifested in various ways'. This presupposes that there is only one substance.<sup>36</sup> That is to say, Spinoza's substance monism resolves several interrelated tensions within Descartes' system noted along the way: those pertaining to the nature of divine concurrence and those



pertaining to the distinction between created and uncreated substance (that is, the distinction gets effaced). Of course, once committed to substance monism, Spinoza can deviate from Descartes, allowing God's 'concurrence' to apply to the 'preservation' of separate (created) entities.

In the second block quote, Spinoza accepts Descartes' diagnosis that there must be some principle which 'connects' an object through different moments of time. But rather than allowing the Cartesian to describe this in terms of constant renewed conservation at each different moment, he prefers the language of constant Divine 'creation as if anew'. On Spinoza's youthful view, the internal preservation of things over time can even be interpreted as constant renewed Divine 'procreation' of things. The language is suggestive because it is a way of talking about the creative aspect of *natura naturans* (a notion already present in *Metaphysical Thoughts* and the *Korte Verhandeling* or *Short Treatise*, I, viii<sup>37</sup>). Moreover, the created entities do not have any (Aristotelian) self-moving principle; they are in an important sense passive (that is, *natura naturata*).<sup>38</sup> But my interest in the second block quote is not exhausted. Spinoza clearly identifies a lacuna in the Cartesian system. Here he accepts Descartes' view that God preserves the same amount of motion in nature. But he notes that there are no Cartesian resources to say anything akin to this about mental phenomena. On Cartesian grounds, it is by no means clear if there is a closed *mental* system. Thus, Spinoza is picking up on the wide domain—covering extended and mental phenomena—of the first law of nature; he has noted Descartes' silence on how to think of its application to the mental economy of nature. So, from Spinoza's point of view, what's missing from Descartes' analysis of internal and (external) 'inertial' preservation of entities is symmetry between extended and mental phenomena. One might say that Spinoza's doctrine of parallelism, enshrined in EIIp7, is designed to solve this problem.<sup>39</sup> Of course, saying this does not make the content of the *conatus* doctrine any less obscure.

#### IV. SPINOZA'S SUBSTANCE MONISM

In the previous sections, I have already indicated and explored some of the tensions internal to Descartes' system. Spinoza's brand of substance monism resolves several such tensions almost by definition. For example, Spinoza's definition of a substance, 'By substance I understand what is in itself and is conceived through itself . . . ' (EID3), suggests another break with Descartes. Recall that for Descartes even the 'supremely perfect' substance (God) is 'the thing in which whatever we perceive (or whatever has objective being in one of our ideas) exists'. While Cartesian substance is self-sufficient, it is not explicitly conceived through itself. Spinoza might well have thought that being conceived through another is a sign of imperfection or lack of proper self-sufficiency.<sup>40</sup> And Spinoza offers an (in-)famous



proof for the claim that 'except God, no substance can be or be conceived' (EIP14).<sup>41</sup>

Thus, if one wishes to resolve Cartesian ambiguities in the direction of simplicity and coherence, then, as we have seen, Spinoza's substance monism is an attractive route: 'Whatever is, is in God, and nothing can be or be conceived without God' (EIP15). Whatever it means that to be is to be in God or be conceived through God, it means that the package of problems surrounding Cartesian Divine concurrence and preservation all must conform to a single template.<sup>42</sup>

Okay, enough detours! Let's return to Spinoza's *conatus* doctrine. Recall EIIIP6:

Each thing naturally strives to persevere in its being.

Dem: [A] For singular things are modes by which God's attributes are expressed in a certain and determinate way (by 1P25C), that is [B] (by IP34) things that express, in a certain and determinate way, God's power, by which God is and acts. [C] And no thing has anything in itself by which it can be destroyed, or which takes its existence away (by P4). [D] On the contrary, it is opposed to everything which can take its existence away (by P5). Therefore, as far as it can, and it lies in itself, it strives to persevere in its being, q.e.d.

Without yet being able to fully resolve all the questions pertaining to the *conatus* doctrine, we are in a position to reject the interpretation of *conatus* as a kind of attribute-neutral external, 'inertial' state-preserving principle (advocated by Vampoulis and Carriero). To be sure, the inertial reading has important advantages in describing human affairs.<sup>43</sup> But as a reading of EIIIP6, it is only plausible if we focus exclusively on the proposition (which does have an inertial-sounding wording) and ignore important aspects of the demonstration. In particular, C-D articulates something far closer to an internal (attribute-neutral) identity preserver of entities than an external (attribute neutral) 'Metaphysical inertia' principle.<sup>44</sup> C&D both make claims about how existence is preserved rather than how states are maintained. (Recall that in Descartes' first law states are preserved.) While the difference may seem trivial (states are also existent), there is no evidence that for Spinoza states are things (note the absence of states at EIIp7S). There is an important metaphysical difference. For an external inertial principle to do its work, an internal entity-preserving principle must already be presupposed.<sup>45</sup> This is, in fact, what C and D (relying on EIIIP4–5) supply. Moreover, as Viljanen has persuasively argued,<sup>46</sup> Spinoza articulates very clear, differently worded (external) inertial principles earlier in the *Ethics* (EIIIP13L3c for bodies).<sup>47</sup> There is no hint that in EIIIP6 Spinoza is relying on them or meant to be restating these.

The whole scholarly discussion in terms of metaphysical inertia has distracted from the meaning of the demonstration. Rather, as Noa Shein



has urged on me, C&D state ‘that insofar as something has being, to that degree it cannot be destroyed—which is just what it means to have being. The real issue is what “things” are there and how much or what kind of “being” so to speak, they have.’ To make this more concrete: the *conatus* doctrine is solving the problem raised by the question: in virtue of what attribute neutral cause do entities remain over time? Nevertheless, Viljanen misleads by speaking of this power in Leibnizian terms as a dynamic resistance to change.<sup>48</sup> For this presupposes that Spinoza is in the same business as Leibniz (and Newton) and is offering us a dynamics that can help us distinguish between real and seeming causes when we measure physical phenomena (something that is very important to Viljanen’s overall interpretation). But even if we ignore the mild anachronism of this approach, there is little reason to think that Spinoza wishes to be read in this way. In Section Vd, I offer an argument against the dynamic reading.

Except for acknowledging Carriero’s point that *conatus* is attribute neutral, which may be implied by A&B, but is best supported by the text of EIIIP9S,<sup>49</sup> I have been very silent on the content of A&B thus far. Moreover, I have said almost nothing thus far about some prominent Spinozistic claims about the nature of *conatus*. For example, in the next proposition, EIIIP7, *conatus* is identified with the actual essence of a thing; and in the proposition after it, Spinoza claims that *conatus* involves ‘an indefinite time’ (EIIIP8). Any satisfactory account of *conatus* must account for these claims. In particular, in order to do justice to A&B, we must become clear on what ‘substance’, ‘attribute’, ‘mode’, ‘affection’, and ‘expression’ mean. So, in what follows I enter into the most contested aspects of Spinoza’s metaphysics. Given the systematic nature of Spinoza’s thought, this is unavoidable.

## V. SUBSTANCE, ATTRIBUTE, MODE, EXPRESSION, AND CONATUS

According to Spinoza, ‘there is nothing except substance and its modes (by A1, D3, D5), and modes (by P25C) are nothing but affections of God’s attributes’ (EIP28Dem). Now in line with Hegel’s influential reading, Wolfson understood the definition of the attributes (EID4) to commit Spinoza to an appearance/reality distinction;<sup>50</sup> this has been contested territory. Hase-rot’s influential response to Wolfson argued that attributes are an ‘objective’ part of Spinoza’s ontology despite the fact that they are not included in the first half of the sentence just quoted.<sup>51</sup> Ever since, the once-common Idealist approach to Spinoza fell out of favour.

Of course, given that modes are defined as ‘affections of a substance’ (EID5), and the existence of modes is unambiguously affirmed (in, for example, the first half of the sentence quoted from EIP28dem), it is tempting to read ‘modes are nothing but affections of God’s attributes’ in an



'objective' fashion. This common temptation runs counter to the fact that the *Ethics*' definition of attribute involves understanding an attribute as 'what the intellect perceives of a substance, as constituting its essence' (EID4).<sup>52</sup> If modes are an official part of Spinoza's 'ontology', then attributes must somehow fit in, too. I take up this challenge, but I prefer to do so by respecting Spinoza's exact wording of the definition of attributes rather than seeing it as badly formulated or misleading. One step forward is to drop the objective versus subjective dichotomy in analysing the nature of attributes.<sup>53</sup> Besides limiting the conceptual possibility space unnecessarily, that dichotomy conflates too easily the language of conceiving (which is used in the definition of substance and mode) with the language of perceiving (which is used in the definition of attribute).

### Va. Finite Modes (1)

Without giving all the details here, a more natural way to read Spinoza's text can be developed if we ask why for Spinoza, given the homogenous nature of substance, that is, all 'matter is everywhere the same' (EIP15S), there should be any 'individual' entities at all. For insofar as 'it is substance, it is neither separated nor divided' (EIP15s; see also Letter 12, On the Infinite, to Meyer, which plays an important background role in my treatment following). To conceive matter in such fashion involves the intellect. It turns out that for Spinoza, only insofar as we conceive imaginatively, are 'parts distinguished' (EIP15s). To distinguish parts within substance involves no (Spinozistic) *real* distinction.<sup>54</sup>

Thus, for Spinoza we conceive of quantity in at least two ways: (a) through our intellect, which sees it without division, as infinite, and homogenous. Spinoza is explicit that this happens only 'with great difficulty' (in the Dutch translation of *Opera Posthuma*). And (b) we conceive of it through the imagination, which allows us to use abstraction in order to 'find' divisible and part-like structures. (EIP15S and Letter 12; for conceiving in Spinoza's system, see Newlands.<sup>55</sup>) To be clear, by imaginative conceiving, one should not immediately think of fiction or falsity (see especially EIIP17s, even though elsewhere [EIIP40S1] Spinoza does have harsh things to say about how imagination and abstraction involve confusion and a lack of adequacy as well as, presumably, lack of completion; see also EIIP16ff).<sup>56</sup>

Now despite its potential unpopularity among those who wish to understand Spinoza as having fairly straightforward views on ontology, this two-ways-of-conceiving-things theory has splendid textual support. In EVP29S, for example, Spinoza writes: 'We conceive things as actual in two ways: either insofar as we conceive them to exist in relation to certain time and place, or insofar as we conceive them to be contained in God and to follow from the necessity of the divine nature.' When we conceive things as existing (determinately) at a certain time and place, we use our imagination or use quantity 'abstractly' to identify their location in space and their



temporal duration as something distinct (at EVP29S Spinoza appeals to EIIP26, but he probably should have also included references to EIP22–25). Thus, according to Spinoza there ‘are’ finite modes if and only if there is imaginative conceiving.<sup>57</sup> That is to say, imaginative conceiving and modes are co-constitutive.<sup>58</sup>

In order to avoid confusion, this means that on my account the third form of knowledge is not about singular things understood as finite modes but, as Spinoza says, about (formal) essences of finite modes (IIP40S2);<sup>59</sup> this knowledge is, thus, of something that singular things have in common (and is cognized by [the] eternal mind[s]). What they have in common is a real definition which when—to use Garber’s language—instantiated affirms the essence of a thing (EIIP4);<sup>60</sup> crucially this is appealed to in EIIP6. Garber is correct to claim that EIIP4–6 follow from what it is to be a nature in Spinoza.<sup>61</sup> That is to say, on my account ‘one’ can ‘perceive’ such formal essences without cognizing (as finite) the finite modes to which they ‘belong’. My interpretation, thus, does not require the introduction of ‘trans-attribute modes’. My approach does retain some ‘attribute-less structure’, that is, formal essences, but these are firmly grounded in Spinoza’s text: EIP17s and, more controversially, IIP40S2. (Of course, there are also formal essences contained in attributes; see EIIP8.)

## Vb. Attributes<sup>62</sup>

By definition, attributes are perceived by the intellect (presumably, by both infinite and finite intellects). Spinoza is clear that attributes are what the intellect perceives of a substance as constituting its essence. I agree with the critics of Wolfson that this is not a statement of subjectivism. I also reject Donagan’s efforts at denying that attributes are genuine essences of the lone substance.<sup>63</sup> Instead, as my comments about the relationship between imagination and finite modes already suggest, I bypass the objective/subjective dichotomy altogether;<sup>64</sup> to state this somewhat awkwardly, intellect and substance are necessarily co-constitutive of the very possibility of the *perception* of attributes. This avoids concerns about attributing to Spinoza extreme versions of (mentalist) idealism.<sup>65</sup> Of course I am avoiding the question of who/what is doing the perceiving!

We are now in a better position to return to details of the *conatus* doctrine. Along the way I explore more about what to make of attributes (as essences of substance, that is, necessary posits for the existence or conceiving of the substance [applying EIID2 back into *Ethics* 1]).

## Vc. Finite Modes and Attributes

Recall the first part of the demonstration of the *conatus* doctrine: ‘[A] For singular things are modes by which God’s attributes are expressed in a certain and determinate way (by 1P25C).’ To be an ‘affection’ of one of God’s



attributes or to be a finite mode means to be 'located' in an infinite chain of (other) finite modes (EIP28). Modes must be conceived through something else (EID5). That is, finite modes, or singular things, 'are nothing but affections of God's attributes' (EIP25C). That is, to say that *what* 'an attribute expresses' or 'to be an affection of an attribute' are two ways of saying the same thing (and this is what EIP25C claims). Fair enough. But EIP25C says something stronger: 'Particular things *are nothing but* affections of God's attributes, or modes by which God's attributes are expressed in a certain and determinate way' (emphasis added).

In the preceding I called attention to the fact that intellect and imagination offer two ways of conceiving 'things'. I claimed that finite modes must presuppose imagistic conceiving.<sup>66</sup> Yet, the demonstration of EIIP6 (and EIP25C) introduces attributes into the equation. And, as we have seen, attributes presuppose/require *intellectual perception*. Even if one were to grant my co-constitutive strategy to resolve some interpretive conundrums, surely imaginative conceiving should not be equated with intellectual perception! So, somebody might be tempted to say that finite modes are ways in which intellect perceives determinate 'affections' of God's essence. I suspect much of contemporary scholarship has tacitly done so without confronting textual difficulties with this move (see my third point following).

We can make progress on this score if we remind ourselves, first, that for Spinoza the intellect falls on the *Natura Naturata* side of things (EIP30–31; so do things that follow from God's attributes; see EIP29S ). In order to perceive an attribute (that is, what the intellect perceives as an essence of God) we must presuppose that it is conceived as an attribute of God. The attribute 'precedes' the intellect because there is no understanding unless there is something to understand with, namely ideas.<sup>67</sup>

Second, if finite beings conceive of individual entities these must always be 'comprehended' (EIIP7s) under an attribute (see also EIP10S). So we can never engage with finite entities directly (even though 'God is really the cause' of 'things as they are in themselves', EIIP7S); this is why they 'are *nothing but* affections of an attribute' (emphasis added). That is, in the demonstration to EIIP6, the parts I labelled 'A and B' articulate considerable conceptual-metaphysical structural apparatus. Spinoza denies that we can conceive individual entities or finite modes directly.<sup>68</sup> In order to be able to speak of determinate entities (or finite modes) at all, we must presuppose three layers of distinctions: imaginative versus intellectual conceptualisation; creative nature versus passive nature; and thinking versus extended attributes.

Third, a finite mode, which is a part of *Natura Naturata* (EIP29S), is a determinate entity in space and time. This way of carving up reality presupposes a number of (modal) distinctions. In particular, it presupposes that we imaginatively conceive of spatial and temporal structures and 'locate' within these a bounded part of extension (and duration) and an infinite chain of causes that can account for the existence of a particular entity



within that region of extension (and duration). Without such imaginative conceptualisation there would be no determinate entities (just substance), and *vice versa*, of course.<sup>69</sup> Given that considerable modal distinctions are presupposed for the conceptualisation of finite modes, it is fair to say that they are not ‘perceived’ directly. Rather, we perceive the same entities either through the attribute of extension, that is, as an affection of extended attribute, or through the attribute of thought, that is, as an affection of thinking attribute (and any other attribute) (EIIP7s).

That is to say, all ‘entities’ are ‘expressions’ of God’s attributes. But while attributes as such are co-constituted by intellectual perception and God, *affections of attributes*, that is, *discerned determinate regions* ‘in’ or ‘expressed by’ an attribute, presuppose applying quantity or other imaginative abstractions (including the application of temporal or spatial structures) to perceived attributes. EIID7 states: ‘By singular things I understand things that are finite and have a determinate existence.’ Spinoza’s definition that singular things are finite and determinate requires that when Spinoza speaks of singular things, the imagination must be involved in conceiving of them.

#### **Vd. God’s Power, Immanent Causation; Formal Causation**

Recall the second part of the *conatus* demonstration: ‘[B] (by IP34) things that express, in a certain and determinate way, God’s power, by which God is and acts.’ God, understood as *Natura Naturans*, ‘really is the cause of things as they are in themselves’ (EIIP7S). But things as they are in themselves are really not finite modes which are never perceived directly. Rather, singular things, being *Natura Naturata*, ‘express’ in a certain and determinate way ‘God’s power’.

Spinoza offers us two doctrines to help clarify what it means for God to be the cause of things as they are in themselves. First, ‘God is the immanent cause of all things’ (EIP18). In the Dutch translation of *Opera Posthuma, Nagelaten Schriften*, the immanent cause is the ‘*inblijvende oorzaak*,’ that is, the cause that stays within. But what does this mean? The accompanying demonstration (EIP18Dem) suggests the doctrine follows from and captures Spinoza’s earlier claim that ‘Whatever is, is in God, and nothing can be or be conceived without God’ (EIP15). So, to say that God is the cause of things (as they are in themselves) involves one in the claim that they cannot be or be conceived without God. Some further clarification on immanent causation is offered by the second doctrine.

Second, according to Spinoza, ‘God is not only the cause of things’ beginning to exist, but also of their *persevering* in existing that is, in scholastic terms, God is the cause of the being of things (*essendi rerum*). For . . . so long as we attend to their essence, we shall find that it involves neither existence nor duration. So their essence can be the cause neither of their existence nor of their duration, but only God, to whose nature alone it pertains



to exist' (EIP24C). This doctrine states that to say that God is the cause of things as they are in themselves is not to speak of their existence in space and time (or the mental economy of nature; see also EVP29S). Rather, it means that God is the (efficient) cause of their being or essence (see also EIP25 and, especially, EIIp45S). So, immanent causation is not about the cause of singular things, that is, finite, determinate entities (located 'in' space and time), but of the essences of things. But whatever essences of things are, they are not, as such, located in space and time. So, on my reading the 'in' part of immanence should not be understood spatially.

EIP24c offers some insight into EIIIP7–8, which any decent interpretation of the *conatus* doctrine must account for anyway. Recall that these identify the *conatus* doctrine with the actual essence of a thing (EIIIP7) and the claim that this striving involves an 'indefinite time' (EIIIP8). I read the latter claim as denying that *conatus* has magnitude. That is, Spinoza is not claiming one can measure *conatus* (as befits his concern over the application of mathematics in EIP15S and L12).<sup>70</sup> So, whatever *conatus* is, it should not (*pace* Viljanen) be viewed akin to a Leibnizian, dynamic magnitude as presupposed in the usual reading of *conatus* as *internal* identity preserver of entities over time. Rather, *conatus* is *essence* preserving.

But what is an individual essence? Now, because of the nod to the Scholastics, several readers have explored the scholastic notion of form (Carriero, Viljanen). A good reason to go along with this is Spinoza's claim that 'from the given essence of each thing some things necessarily follow (by EIP36)' (EIIp7dem). That is, an essence is clearly a certain kind of cause. Moreover, knowing an essence means one possesses a so-called real definition (see especially EIP33S1, but also, of course, EIIIP4dem and letter IX to De Vries, TIE §50–52). Nevertheless, I agree with Laerke that this is a nonstarter because the Scholastic notion presupposes hylomorphism. If one goes the 'form' route, Karolina Hubner's arguments (building on Mancosu's study) about the Platonic-mathematical notion of form are very promising.<sup>71</sup>

Nevertheless, here I argue that Spinoza's account is similar to a Baconian notion of form.<sup>72</sup> Consider the following passage:

On a given body, to generate and superinduce a new nature or new natures is the work and aim of human power. Of a given nature to discover the form, or true specific difference, or nature-engendering nature, or source of emanation (for these are the terms which come nearest to a description of the thing), is the work and aim of human knowledge. Subordinate to these primary works are two others that are secondary and of inferior mark: to the former, the transformation of concrete bodies, so far as this is possible; to the latter, the discovery, in every case of generation and motion, of the *latent process* carried on from the manifest efficient and the manifest material to the form which is engendered; and in like manner the discovery of the *latent configuration* of bodies at rest and not in motion.



In what an ill condition human knowledge is at the present time is apparent even from the commonly received maxims. It is a correct position that “true knowledge is knowledge by causes.” And causes again are not improperly distributed into four kinds: the material, the formal, the efficient, and the final. But of these the final cause rather corrupts than advances the sciences, except such as have to do with human action. The discovery of the formal is despaired of. The efficient and the material (as they are investigated and received, that is, as remote causes, without reference to the latent process leading to the form) are but slight and superficial, and contribute little, if anything, to true and active science. Nor have I forgotten that in a former passage I noted and corrected as an error of the human mind the opinion that forms give existence. For though in nature nothing really exists besides individual bodies, performing pure individual acts according to a fixed law, yet in philosophy this very law, and the investigation, discovery, and explanation of it, is the foundation as well of knowledge as of operation. And it is this law with its clauses that I mean when I speak of *forms*, a name which I the rather adopt because it has grown into use and become familiar (*New Organon*, 2.I–II; emphasis in original).<sup>73</sup>

Bacon is no Scholastic, as can be clearly seen by his claims that (a) ‘final cause rather corrupts than advances the sciences, except such as have to do with human action’, and (b) ‘that forms give existence’ is ‘an error’. Yet, Bacon embraces the three other Aristotelian causes (material, formal, and efficient) as entirely appropriate to his proposed science of nature. In particular, here I focus on Bacon’s willingness to embrace a reformulated version of formal causation; ‘the aim of human knowledge’ is to discover the ‘nature-engendering nature’ of *bodies* in order to create new ‘natures’. Bacon’s forms are materialistic. In labelling his causes in this fashion, Bacon is even willing to risk confusion with the Scholastics.

Now, despite their shared hostility to final causes in physics and their shared interest in natural history, Spinoza and Bacon do have serious disagreements; there is no denying that Spinoza is no Baconian. Spinoza’s metaphysics is incompatible with Bacon’s foundational claim that ‘in nature nothing really exists besides individual bodies’. Nevertheless, Spinoza explicitly echoes Bacon’s notion (and language) of *form* in Lemmas 4–7 in (to use Lachterman’s apt phrase) the physical interlude between EIIp13 and p14. Exploring that fully must be left for another time. Here I merely insist that we can read Spinoza’s *conatus* as generalizing Bacon’s ‘nature-engendering nature’. Such a nature-engendering nature conforms rather nicely to Spinozistic formal essences of things. In Bacon an essence is exclusively materialistic; in Spinoza that would be the actual essence of a thing as conceived under the attribute of extension. However, the essence of a thing, which *conatus* preserves, is ‘attribute-neutral’. A mode must be conceived through an attribute. But conceiving of an individual essence



does not require that the essence be conceived through an attribute. It is only when the individual essence is 'related' to a mode that it has to be conceived through an attribute. Of course, when it is 'related only to the mind, it is called will; but when related to the mind and body together, it is called appetite' (EIIIp9S).<sup>74</sup> As a 'non-goal-directed' (formal) cause it determines the thing of which it is an essence.<sup>75</sup> The *conatus* is God's presence in all things. Thus, *conatus* is an attribute-neutral essence preserving immanent cause.<sup>76</sup> Of course, this phrase cannot substitute for explaining the way of being and our access to such a cause.

Moreover, it follows from the reading developed here that the striving and reality of finite things is in some important sense illusory. It also means that Spinoza has not gotten very far from Descartes' position; talk of an individual nature is a way of talking about God's activity. Even so, in Spinoza's system God's activity presupposes, in some nontrivial sense, finite beings like us and this sets him apart from Descartes. But explaining such mysteries must be left for another occasion.

## NOTES

1. I thank Bruno Verbeek, Marietje van Schaar, Bart Karstens, and my students, Dagmar Hepp, Jan van der Ham, and Guus van Zwoll, who helped me think through a lot of these issues in a reading group. I also received very helpful, detailed, and very critical comments from Charlie Huenemann, Valtteri Viljanen, Sam Newlands, Andrew Youpa, Noa Shein, and Mogens Laerke on earlier drafts of this chapter. Special thanks are due to the audience members at Syracuse University's Spawn 2009 event and my insightful commentator, Michael Della Rocca.
2. I am not committed to three stronger theses: (a) that Spinoza's *conatus* doctrine was (solely) intended as a response to these problems; (b) that Spinoza's *conatus* doctrine can only be understood if one understands what problems gave rise to it; (c) that for Spinoza there are no other problems in Descartes (see, e.g., Letter 81, 5 May 1676, and Letter 83, 15 July 1676, both to Tschirnhaus).
3. In an early Letter to Oldenburg, 17/27 July, 1663, Spinoza makes clear that he hopes that his commentary on Descartes' *Principles* will attract attention of well-placed dignitaries, so he can receive their protection to publish his own views.
4. Gary Hatfield, "Reason, Nature, and God in Descartes", in *Essays on the Philosophy and Science of René Descartes*, edited by Stephen Voss (Oxford: Oxford University Press, 1993); Catherine Wilson, *Epicureanism at the Origins of Modernity* (Oxford: Oxford University Press, 2008).
5. I have been unable to find a place where Spinoza calls anything but God an 'immanent' cause (see also Letter 73 to Oldenburg, December 1675). I hope my reasons for using 'immanent' are clear near the end of the chapter.
6. I refer to Spinoza's *Ethics* by part in roman numeral and then the proposition (and if necessary its demonstration, scholium, or corollary). Unless otherwise noted, my translations follow Edwin Curley's in *A Spinoza Reader: The Ethics and Other Works* (Princeton: Princeton University Press, 1994).
7. Jonathan Bennett, *A Study of Spinoza's Ethics* (Cambridge: Cambridge University Press, 1994).



8. Edwin Curley, *Behind the Geometrical Method: A Reading of Spinoza's Ethics* (Princeton, NJ: Princeton University Press, 1988); Martin Lin, "Spinoza's Metaphysics of Desire: The Demonstration of IIP6", *Archiv für Geschichte der Philosophie*, 86 (2004): 21–55; Don Garrett, "Spinoza's Conatus Argument", in *Spinoza: Metaphysical Themes*, edited by O. Koistinen & J. Biro (Oxford: Oxford University Press, 2002); and especially, Valtteri Viljanen, "On the Derivation and Meaning of Spinoza's Conatus Doctrine", *Oxford Studies in Early Modern Philosophy*, 4 (2008): 89–112.
9. Henry Austryn Wolfson, *The Philosophy of Spinoza: Unfolding the Latent Processes of His Reasoning* (Cambridge: Harvard University Press, 1934).
10. Paul Oskar Kristeller, "Stoic and Neoplatonic Sources of Spinoza's *Ethics*", *History of European Ideas*, 5 (1984): 1–15.
11. Marin Lin, *Spinoza's Theory of Desire*, dissertation defended at the University of Chicago (2001).
12. Daniel Garber, "Descartes and Spinoza on Persistence and Conatus", *Studia Spinozana*, 10 (1994): 43–67.
13. Michael Della Rocca, "Spinoza's Metaphysical Psychology", in *The Cambridge Companion to Spinoza*, edited by D. Garrett (Cambridge: Cambridge University Press, 1996).
14. Tammy Nyden, *Spinoza's Radical Cartesian Mind* (London: Continuum, 2007).
15. David Lachterman, "The Physics of Spinoza's *Ethics*", in *Spinoza: New Perspectives*, edited by R. W. Shahan & J. I. Biro (Norman: University of Oklahoma Press, 1978); Garber, "Descartes and Spinoza on Persistence and Conatus".
16. Here I amend Curley in light of I. Bernard Cohen, "'Quantum in Se Est': Newton's Concept of Inertia in Relation to Descartes and Lucretius", *Notes and Records of the Royal Society of London*, 19 (1964): 131–155.
17. One can say this while remaining agnostic as to what degree Spinoza intends or manages to banish all forms of teleology from his system.
18. Of course, Don Garrett, "Teleology in Spinoza and Early Modern Rationalism", in *New Essays on the Rationalists*, edited by R.J. Gennaro & C. Huenemann (Oxford: Oxford University Press, 1999), thinks this is great evidence for a teleological reading of Spinoza.
19. EIID2: "I consider as belonging to the essence of a thing that, which being given, the thing is necessarily given also, and, which being removed, the thing is necessarily removed also; in other words, that without which the thing, can neither be nor be conceived."
20. I quote Descartes' *Principles of Philosophy* with minor modifications from René Descartes, *Principles of Philosophy*, translated by V. R. Miller and R. P. Miller (Dordrecht: Kluwer, 1983). Other references are to R. Descartes, *The Philosophical Writings of Descartes, Volume II*, translated by John Cottingham et al. (Cambridge: Cambridge University Press, 1984).
21. Perhaps God also has a principal attribute, existence? (If so, it is worth exploring if existence is one of the attributes in Spinoza.) Igor Agostini has called my attention to Descartes' *Principles of Philosophy* part I, art. LVI (AT VIII-1 26, ll. 27–30): '... *etiam in rebus creatis . . . existentia et duratio . . . attributa dici debent*', but it does not claim existence is the principal attribute.
22. See Tad Schmaltz, "Descartes on the Extensions of Space and Time", *Revista Analytica* (forthcoming), and Alison Simmons, "Comments on Tad Schmaltz's 'Descartes on the Extensions of Space and Time'" (presented at NYU Conference, 10 November 2006), for contrasting views. See also Tad Schmaltz, *Descartes on Causation* (Oxford: Oxford University Press, 2008),



- and Andrew Pessin, "Descartes' Nomic Concurrentism: Finite Causation and Divine Concurrence", *Journal of the History of Philosophy*, 41 (2003): 25–49. Also note, by the way, I do not mean to imply that for Descartes time is substantial. The partition of time is used to motivate a separability thesis.
23. Note that the way I am using 'separability thesis' I mean that separability equals to possible independent existence; independent conceivability relates to a real distinction.
  24. Garber, "Descartes and Spinoza on Persistence and Conatus", 52, claims it follows from God's providence. If so, this would motivate Spinoza's dissent.
  25. This is subject of controversy, but see Helen Hattab, "The Problem of Secondary Causation in Descartes: A Response to Des Chene", *Perspectives on Science*, 8 (2000): 93–118.
  26. One might read Descartes' use of the phrase '*quantum in se est*' (*Principles*, II.37) to describe this. See Cohen's analysis of the phrase (in "'Quantum in Se Est'") as claiming that to Descartes' seventeenth-century readers it would indicate what a thing does 'by nature' or 'naturally' (as I have chosen in my translation of Spinoza). If we accept Cohen's reading, then Descartes is explicitly assuming that entities have an internal identity-preserving cause, but the question of its nature would still remain. I thank Della Rocca for pressing this line.
  27. While I tend to downplay Newton's debts to Descartes, this methodological understanding of the laws of nature is a very prominent feature of Newton's use of the laws of motion.
  28. In a letter to De Volder, Leibniz captures a part of the problem: 'It is one thing to retain a state until something changes it, which even something intrinsically indifferent to both does, and quite another thing, much more significant, for a thing not to be indifferent, but to have a force and, as it were, an inclination to retain its state, and so resist changing', quoted by Garber, "Descartes and Spinoza on Persistence and Conatus", 47. Garber's treatment of problems with Descartes' laws of motion is exemplary.
  29. See, e.g., *Principles of Philosophy*, 2.42.
  30. In his exchanges with Hartsoeker published in February 1711, Leibniz appealed to 'conspiring motions' to account for this. See Isaac Newton, *Philosophical Writings*, edited by A. Janiak (Cambridge: Cambridge University Press, 2004), 110.
  31. This is not to deny that Descartes could tell a story involving all the beneficial effects of God's permanence. Or, as Michael Della Rocca reminded me, Descartes could simply deny that there is mental inertia; he need not be a naturalist.
  32. Lachterman, "The Physics of Spinoza's *Ethics*", 99, which draws on Descartes' *Principles of Philosophy* 2.43 and 2.55.
  33. To be sure: this should not be understood as an argument against explanatory reductionism.
  34. Garber, "Descartes and Spinoza on Persistence and Conatus"; Épaminondas Vampoulis, "Le principe d'inertie et le conatus du corps", *Astérion*, 3 (2005): <http://asterion.revues.org/document304.html> [note that the journal spells Vampoulis's name mistakenly as 'Vamboulis']; John Carrier, "Spinoza on Final Causality", *Oxford Studies in Early Modern Philosophy* II, edited by D. Garber & S. Nadler (Oxford: Oxford University Press, 2005).
  35. Michael Della Rocca, *Spinoza* (London: Routledge, 2008).
  36. For useful caution on the nature of Spinoza's 'monism', see P. Macherey, "Spinoza est-il moniste?" *Spinoza. Puissance et ontologie*, edited by M. R. D'Allones and H. Rizk (Paris: Kimé, 1994), 39–53, and Mogens Laerke, *Leibniz lecteur de Spinoza* (Paris: Champion, 2008), especially 671–78.



37. Benedict Spinoza, *Korte Geschriften*, edited by F. Akkerman et al. (Amsterdam: Wereldbibliotheek, 1982).
38. As Noa Schein has pointed out to me, they are not finite on their own accord but determined by all the other finite things. They need not, of course, be ‘passive’ in the sense that matter is passive in, say, Descartes’ philosophy.
39. In private correspondence, Charlie Huenemann speculates that Descartes’ voluntarism may be the source of the problems.
40. Interestingly, only *natura naturans* is conceived through itself, while *natura naturata* conceives of nature through the things that follow from God’s necessity, that is, the modes, as ‘things which are in God and can neither be nor be conceived without God’ (1Ep29S).
41. Some might think that for Spinoza attributes are also conceived through themselves; my interpretation of attributes below blocks this move. I thank Valtteri Viljanen for raising this issue.
42. Cf. O. Koistinen, “Causation in Spinoza”, *Spinoza: Metaphysical Themes*, edited by O. Koistinen & J. Biro (Oxford: Oxford University Press, 2002).
43. See Carriero, “Spinoza on Final Causality”, 139, and Valtteri Viljanen, *Spinoza’s Dynamics of Being: The Concept of Power and Its Role in Spinoza’s Metaphysics*, dissertation defended at University of Turku (2007), 177.
44. Carriero, “Spinoza on Final Causality”, 134.
45. There are three issues lurking here: (i) a difference that makes a difference needs to be explained; (ii) persistence is not primitive in Spinoza’s system; (iii) it is quite possible (as my view entails) that no such entity-preserving principle can be found spatially inside a thing—Noa Shein, for example, has argued that modes co-determine.
46. *Spinoza’s Dynamics of Being*, 181–82.
47. From this it follows that a body in motion moves until it is determined by another body to rest; and that a body at rest also remains at rest until it is determined to motion by another (II3p13L3C).
48. *Spinoza’s Dynamics of Being*, 182ff. In Leibniz, velocity and direction equal conatus/nisus, so that in Leibniz this is all very teleological. I thank Mogens Laerke for helpful discussion.
49. ‘When the striving is related only to the mind, it is called will; but when it is related to the mind and body together, it is called appetite. This appetite, therefore, is nothing but the very essence of man, from whose nature there necessarily follow those things that promote his preservation. And so man is determined to do those things. . . .’
50. *The Philosophy of Spinoza*, 146–56.
51. Francis S. Haserot, “Spinoza’s Definition of Attribute”, *The Philosophical Review*, 62 (1953): 499–513.
52. Newton scholars will not be surprised to learn that Spinoza’s *tanquam is the nub*. See Alexandre Koyre and I. Bernard Cohen, “The Case of the Missing Tanquam: Leibniz, Newton & Clarke”, *Isis*, 52 (1961): 555–66.
53. Noa Shein, “The False Dichotomy between Objective and Subjective Interpretations of Spinoza’s Theory of Attributes”, *British Journal for the History of Philosophy*, 17 (2009): 505–32, and “Spinoza’s Theory of Attributes”, *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, <http://plato.stanford.edu/archives/spr2009/entries/spinoza-attributes/>, has made thorough progress on this score. We developed our interpretations independently from each other, although we have had inspirational conversations on these matters. On some key points our approach is structurally similar. But she ends up solving interpretive problems through a Cartesian-inspired epistemic reading of Spinoza’s doctrine of real distinctions. My approach focuses on ontology. It appears our solutions cohere with and may mutually support



each other. Mogens Laerke points out that my reading is very much in the spirit of Schelling's reading of Spinoza in his seminars during the 1830s. Schelling and I are both struck by the fact that change and (what one may call) limitation seem at first glance at odds with the nature of Spinozistic substance.

54. There are issues lurking here about what kind of intellects can perceive finite entities at all.
55. Samuel Newlands, "Another Kind of Spinozistic Monism", *Nous* (forthcoming).
56. In what follows I am going to pretend that there are no problems with Spinoza's use of 'quantity' (*quantitatis*) here; because at EIIP45S it appears that Spinoza identifies quantity with the process of abstraction and this is, I think, the fully authentic Spinozistic position. I treat Spinoza's undivided quantity (of EIP15S) as a way of speaking about pure extension.
57. Noa Shein has argued for the foundational significance of EII, Axiom 4: 'We feel that a certain body is affected in many ways.' This would allow us to affirm both sides of the biconditional. (Not a bad result for a transcendental argument!) But I cannot explore that here.
58. One reader proposed that 'from the standpoint of the intellect, there are finite modes, but they are not regarded as having a specific spatial and temporal location'; but if such regarding were possible it would be better to say that 'they' are, thus, not regarded as modes. For important comments see Shein, "The False Dichotomy between Objective and Subjective Interpretations of Spinoza's Theory of Attributes" and "Spinoza's Theory of Attributes", 1.9.4.
59. Controversially, I am accepting the Dutch translation of *Opera Posthuma* as authoritative here.
60. How this is supposed to work is by no means obvious. I articulate Spinoza's views in Eric Schliesser, "Spinoza and Science", *Oxford Handbook of Spinoza*, edited by Michael Della Rocca (Oxford: Oxford University Press, forthcoming).
61. "Descartes and Spinoza on Persistence and Conatus", 62.
62. For a terrific introduction, see Shein, "Spinoza's Theory of Attributes".
63. Alan Donagan, *Philosophers in Context: Spinoza* (Hertfordshire: Harvester Wheatsheaf, 1988).
64. See also Shein, "The False Dichotomy between Objective and Subjective Interpretations of Spinoza's Theory of Attributes".
65. EIIP7S is very useful: 'The thinking substance and extended substance are one and the same substance, which is now comprehended under this attribute, now under that. So also a mode of extension and the idea of that mode are one and the same thing, but expressed in two ways . . . So of things as they are in themselves, God is really the cause insofar as he consists of infinite attributes' (emphasis added). Spinoza comes tantalisingly close to asserting that things in themselves are inaccessible to finite intellects which cannot comprehend under multiple attributes simultaneously (perhaps because of what Della Rocca, "Spinoza's Metaphysical Psychology", has labelled the 'conceptual barrier' between the attributes).
66. There are concerns over (numerical) identity between 'things' as conceived by intellect and imagination, but I cannot explore these in detail here; I thank Sam Newlands for revealing conversation on these matters.
67. Note that this introduces an element of idealism into Spinoza's system!
68. This claim should not be confused with the well-known Spinozistic claim that we perceive entities through affections of our own bodies.
69. Here I am agnostic about the status of infinite modes.



70. Much damage has been done by scholars who want to make Spinoza part of the main thrust of mechanical philosophy (broadly conceived). For example, Garber, “Descartes and Spinoza on Persistence and Conatus”, 64, writes, ‘Spinoza wants to explain everything in terms of size, shape, and motion’. In Schliesser, “Spinoza and Science”, I argue that Spinoza’s conceptual categories are far removed from this project. Spinoza would deny that employing the categories ‘size, shape and motion’ could ever provide stable knowledge. For all of Spinoza’s understanding of and participation in contemporary natural philosophy, Spinoza’s project is not aimed at knowledge of nature in the sense (however different) of Descartes or Bacon. (Spinozistic knowledge is really a form of self-knowledge, even though the eternal self is then dissolved.) The letter to Bouwmeester, 10 June 1666, on the source of clear and distinct ideas is very useful on these matters.
71. K. Hubner (ms), “On the Significance of Formal Cause in Spinoza’s Metaphysics”; Paolo Mancosu, *Philosophy of Mathematics and Mathematical Practice in the Seventeenth Century* (Oxford: Oxford University Press, 1999).
72. In his correspondence with Oldenburg, Spinoza makes a point of asserting his familiarity with Bacon. Alan Gabbey, “Spinoza’s Natural Science and Methodology”, in *The Cambridge Companion to Spinoza*, edited by D. Garrett (Cambridge: Cambridge University Press, 1996), offers useful cautionary comments.
73. Francis Bacon, *The New Organon: Or True Directions Concerning the Interpretation of Nature*, in *The Works*, Vol. VIII, translated by James Spedding et al. (Boston: Taggard and Thompson, 1863[1620]): [http://www.constitution.org/bacon/nov\\_org.htm](http://www.constitution.org/bacon/nov_org.htm).
74. On my reading, the will and appetite presuppose imaginative conceiving.
75. If the Dutch translation can be trusted, Spinoza also has claims about ‘formal essence of things’ (EIIP40S2), but it is by no means obvious these ‘things’ are genuinely individuated (see Spinoza’s own warning at EIIP45S).
76. That is, *conatus* is in Della Rocca’s terms a *neutral property*; but in the application of this concept I disagree with Della Rocca (see notes 54 and 58). On my reading, there is no attribute-neutral mode that is a bearer of neutral properties.



# 5 Are Mind-Body Relations Natural and Intelligible? Some Early Modern Perspectives

*Pauline Phemister*

Between 1703 and 1705, Damaris Masham and Leibniz exchanged a series of letters, parts of which touched briefly on the relative merits of Cartesian interactionism, Malebranchean occasionalism and Leibnizian pre-established harmony.<sup>1</sup> Despite the brevity of Leibniz's defence of pre-established harmony against Masham's criticism of both it and occasionalism, the exchange, considered alongside other texts from the same period, sheds light on some of the underlying reasons for Leibniz's advocacy of pre-established harmony over the other two theories.

Masham's own preference for Cartesian interactionism rests on her belief that, of the three theories, interactionism is the one that most explicitly requires, and therefore justifies, belief in the structured organisation of the body. She charges occasionalism and pre-established harmony with the same fault: neither explains why God created bodies possessing the intricate and ordered constructions that microscopic investigations had recently revealed. Under occasionalism and pre-established harmony, she contends, 'all that nice curiositie that is discoverable seeming useless, becomes superfluous and lost labour' (Masham to Leibniz, 3 June 1704: GP III.350; WF 209). The observation is perspicacious. After all, according to the doctrine of occasionalism, minds' sensations of bodies are caused directly by God, while by the doctrine of pre-established harmony, sense perceptions corresponding to body states emerge naturally from the inner depths of the soul itself. On neither view does the intricate organisation of bodies' innermost parts play a causal role in the soul's perception of bodies' outwardly visible and sensible qualities. Neither doctrine would seem therefore to offer a reason why bodies need to have intricately organised parts at all. The organisation of bodies is an unnecessary complication for a God who presumably acts with parsimony in accordance with the principle of economy. In contrast, according to the doctrine of interactionism, the internal mechanical structures of bodies are essential if external bodies are to act on our own in



such as way as to effect changes to our brains from which result the ideas and perceptions in our souls.

All the same, exactly how physical impressions imposed on the brain are transformed into ideas and perceptions in the mind is obscure. Both Malebranche and Leibniz were quick to note interactionism's failure to provide an exact account of precisely *how* mind and body act on each other. Malebranche rejected interaction because he considered that interaction would require that the soul have knowledge of the motion of the insensible parts of its body—and this, he believed, it does not have. Leibniz rejected interactionism on the ground that interaction between substances with totally different attributes is impossible and because the relation between mind and body postulated by Descartes is both the result of an arbitrary divine volition and consequently, in his opinion, humanly unintelligible.

Descartes had actually regarded interaction between mind and body as more intelligible (or less unintelligible) than the actions of heat or weight on a body, for the former postulates the interaction between substances while the latter postulates interaction only between an accident and a substance. The problem, he diagnosed,

arises simply from a supposition that is false and cannot in any way be proved, namely that, if the soul and the body are two substances whose nature is different, this prevents them from being able to act on each other. And yet, those who admit the existence of real accidents like heat, weight and so on, have no doubt that these accidents can act on the body; but there is much more of a difference between them and it, i.e. between accidents and a substance, than there is between two substances. (Appendix to *Fifth Set of Replies*: AT IXa.213, CSM II 275)

Nonetheless, Cartesian interaction between the substances, mind and body, remains in principle unintelligible on account of its being the result of an arbitrary decree of a voluntaristic God. Defending interactionism, Masham simply refused to grant that God must act in ways that are intelligible to human understanding. 'Gods ways', she wrote to Leibniz on 3 June 1704, 'are not limited by our conceptions; the unintelligibleness or inconceivableness by us of any way but one, dos not methinks, much induce a Beleeve of that, being the way which God has chosen to make use of' (GP. III.350, WF 209).

Leibniz wholly disagrees. For him, the way God arranges relations between minds and bodies must be intelligible not only to God, but also to humans. Accordingly, in his next letter dated 30 June 1704, he defended the convergence of human and divine intelligence, pointing to recent astronomical and physical studies as empirical support for direct proportionality between the intelligibility of a theory and the likelihood of its being true.<sup>2</sup> He refused to admit the gulf between God and human reason implied by Masham's defence of interactionism. Instead, he insisted upon a strong



correlation between a theory's truth and its human comprehensibility, and supported his claim by appeal to the rationality of the divine source of human understandings. That our understandings derive from God vindicates our belief that when we employ our understandings well and reason methodically in accordance with the divinely granted natural light, our conclusions conform to God's understanding and wisdom (GP.III.353, WF 211). However, the divinely established but thoroughly arbitrary relation between sensations and bodily motions proposed by the interactionists is beyond rational comprehension. No matter how well we employ our understandings or reason methodically, we will never be able to fathom the causal connection between mind and body.

Leibniz rejected occasionalism for the same reason. Malebranche had accepted the arbitrary and unintelligible nature of psycho-physical interaction, but instead of searching, as Leibniz believed they should, for a humanly intelligible account of mind-body relations, the occasionalists had invoked the miraculous and unintelligible operation of a *deus ex machina* (*New System of the Nature and Communication of Substances*: D II.54, WF 17). This 1695 remark was repeated in the correspondence with Masham: occasional causation requires miraculous action, the understanding of which lies outside the range of human intellect (Leibniz to Masham, 30 June 1704: GP III.354, WF 212). Only pre-established harmony explains the relation of mind and body by natural, non-miraculous secondary causes that are intelligible to both God and humans. Thinking that interactionism, occasionalism, and pre-established harmony exhaust the possible options, and supposing that 'ordinary things must happen naturally and not by a miracle', Leibniz declared his theory demonstrated (Leibniz to Masham, 30 June 1704: GP III.355, WF 213).<sup>3</sup>

Leibniz's defence of the human intelligibility of mind-body relations not only allowed him to fend off the opposition; it also afforded him the means to respond to Masham's challenge (3 June 1704: GP III.350, WF 209) that he justify his claim in his letter of early May 1704 that 'organism, that is to say, order and ingenuity, is something essential to matter produced and arranged by the sovereign wisdom' (Leibniz to Masham, early May 1704: GP III.340, WF 205). His answer posits the organisation of the body as essential if mind-body relations are to be natural and non-miraculous and thereby in principle intelligible to human understanding. In the following sections, we explore his response more fully.

## I

Natural relations are distinguished from miraculous relations by reference to the ways in which God acts (Leibniz to Masham, 30 June 1704: GP III. 353, WF 211). The former result from God's acting by natural means and are in principle comprehensible to human minds, given sufficient



opportunity and guidance to investigate their causes. The latter arise when God acts in miraculous ways and are such that, no matter how much attention is given them, understanding of them will always lie outside the reach of human comprehension.

For Leibniz, ‘immediate communication’ between ‘heterogeneous substances’, such as that postulated by Descartes’ interactionism, must be miraculous (Leibniz to Masham, 30 June 1704: GP III.354, WF 212). Malebranchian occasional causes also postulate immediate communication between heterogeneous substances, God and created beings,<sup>4</sup> and are equally miraculous and unintelligible. In Leibniz’s view, the relation of mind and body proposed by Malebranche depends ‘solely on the immediate operation of God, and allows of no other means or explanation’ (ibid.). Malebranche’s God must act not only as general cause insofar as He determines the general laws that govern psycho-physical relations, but also as particular cause because Malebranche refused to grant causal powers to particular things:

It isn’t sufficient for an action not to be miraculous that it be in conformity with a general law. For if that law had no foundation in the nature of things, perpetual miracles would be needed to implement it . . . Thus it is not enough for God to ordain that the body should obey the soul, and that the soul should have perceptions of what happens in the body; he must also give them some means of doing so. (Comments on Bayle’s *Historical and Critical Dictionary*, ‘Rorarius’ Note H: GP IV. 533, WF 78)

Malebranche denied minds and bodies their own internal active forces: there are no *natural* immediate causes of particular changes in minds and bodies. Leibniz surmised therefore that, since minds and bodies are incapable of effecting the changes themselves, each occasional cause must require God’s miraculous action. By Leibniz’s reckoning, for the occasionalists, whether they admit it or not (to Masham, 30 June 1704: GP III.354, WF 212), the immediate cause of each particular mind-body correlation is supernatural, not natural.

But what is the connection between the sensations in souls and the configurations of particles in bodies that comprise their essences? How are the sensible perceptions of bodies related to physical reality? How are sensory perceptions of colours and sounds correlated with the motion and construction of bodies? Malebranche’s uncompromising answer is that there is ‘no relation or necessary connection between the disturbances of the brain and particular sensations of the soul’ (*Dialogues on Metaphysics and on Religion*, Dialogue 4: OC XII.93; JS 56). God has established general psycho-physical laws that maintain a regular and invariable correlation between various types of motion of the extended parts of bodies and types of sensations in our souls. But that *this* type of configuration of particles in bodies



is correlated with, for instance, a yellow sensation in the soul is a result of 'laws of the union of the soul and body' that, as Malebranche makes clear in the *Dialogues on Metaphysics and on Religion*, are entirely *arbitrary* (Dialogue 6: OC XII.138, JS 96, my emphasis). Leibniz read Descartes as taking the same line, reporting in the *New Essays* that 'the Cartesians . . . regard it as arbitrary what perceptions we have of these qualities, as if God had given them to the soul according to his good pleasure, without concern for any essential relation between perceptions and their objects' (Preface: A VI vi 55, RB 55).

As noted earlier, Leibniz considered such non-essential arbitrary correlations between souls' sensations and bodily states as humanly incomprehensible. After all, there is no necessity that certain types of sensations be correlated with certain types of bodily states rather than with others. The God of Descartes and Malebranche could presumably have correlated different types of sensations with bodily motions, perhaps by reversing the colour spectrum or consistently correlating souls' aural sensations to those configurations of our brains that are in this world correlated with visual sensations. We can discover by empirical investigation which types of sensations are correlated with which types of bodily states. However, the reason why these particular correlations rather than others obtain is unintelligible to humans and quite possibly also unintelligible even to God, for they are neither divinely grounded in God's rational choice nor naturally grounded in the essences, powers or natures of souls and bodies.<sup>5</sup>

Leibniz's own view is that mind-body relations are necessary rather than arbitrary and also in principle humanly comprehensible. They are also natural. From their intelligibility, he concluded that the relations are explicable solely in terms of the laws of nature: 'what is comprehensible is what conforms to the natural laws of things, and should be explained only by them' (Leibniz to Masham, 30 June 1704: GP III.354, WF 212). According to the system of pre-established harmony, the succession of perceptions and appetitions, desires or volitions in souls and the movement of bodies from one place to another are explicable in terms of natural laws that govern the activities of these two very different types of entities. Souls are subject to the laws of final causes that employ moral and aesthetic concepts such as goodness, right, perfection, beauty, virtue, love, and justice. Bodies, on the other hand, are subject to the laws of efficient causes that employ concepts from the mechanical philosophy such as force, motion, extension, mass, figure, resistance, inertia, impetus, and velocity (*Monadology*, §79: GP VI.620, AG 223).

The changes within the sequences of perceptions in souls and movements in bodies not only accord with the general laws of nature; they are also effected by minds and bodies themselves. The movement from one perception to the next is brought about by the soul's own appetite, that is, by the action of the primitive force that characterises the soul's individual



essence. Similarly, the movement of a body from one location to another is brought about by the body's own derivative or physical force.<sup>6</sup> Thus, the immediate causes of changes in created things are natural, being due to forces internal to the souls and bodies themselves. And natural causes are in principle intelligible to human beings who can, through reason and observation, discover the laws that govern the realms of grace (the laws of final causes) and of nature (the laws of efficient causes), respectively.

All the same, it is not obvious that Leibniz's system of pre-established harmony fares any better than Malebranche's doctrine of occasionalism in providing an intelligible account of relations between minds and bodies. Even if causation internal to each of the respective series of efficient causes in bodies and final causes in souls is open to human comprehension so that we can understand how the motion of one body brings about motion in another body and how the soul's present perceptions are related to its past and future perceptions, the vexed issue of the intelligibility of the mind-body relation, of the correspondence between motions in the body and perceptions in the soul, remains. Leibniz has asserted only that God acts by natural means in respect of each sequence (of perceptions and of motions) conceived separately. He has not, so far at least, explained why the correspondence between these two sequences is not itself merely arbitrarily imposed by God. The ordered relation between Leibnizian souls' perceptions and the motions of the parts of bodies might consist only of non-essential relations, imposed simply by God acting by 'miraculous ways' that are neither natural nor humanly intelligible. If so, then the relation between souls' sensations and bodily motions would be just as inexplicable on Leibniz's hypothesis as it is on Malebranche's.

## II

Before addressing this problem directly, it will be helpful to pick apart two separate issues that Leibniz conflated. On the one side, there is the question whether the correspondence between perceptions and motions is rational, essential, and intrinsically humanly intelligible or purely arbitrary and unintelligible. On the other side is the question whether God brings about the correspondence by natural or by miraculous means.

It is clear that Leibniz considered these two questions to be inextricably linked. For him, whatever occurs by natural means, in accordance with laws of nature, is reasonable, essential, and humanly intelligible; whatever God brings about by miraculous means is arbitrary and humanly unintelligible. But other combinations are possible. Mental and physical states may be related to each other in ways that are:

- (i) essential and intelligible, and brought about by God acting by natural means



- (ii) essential and intelligible, but brought about by God acting miraculously
- (iii) arbitrary and unintelligible, but brought about by God acting by natural means
- (iv) arbitrary and unintelligible, and brought about by God acting miraculously

Leibniz's pre-established harmony is captured by (i) and Malebranche's occasionalism by (iv). Spinoza's account of mind-body relations might be read as a version of (ii), albeit with one reservation which I mention below. For Spinoza, the changing states of bodies are necessary and intelligible in terms of physical causation and are paralleled by the logically necessary and equally intelligible connections among the mind's ideas (*Ethics* II, Prop. 7). However, Spinoza viewed minds and bodies as composite modes arising respectively from God's attributes of thought and extension. They follow necessarily from the *divine* nature, as does also their correspondence insofar as both idea-modes and body-modes are modes of attributes that express the same divine essence.<sup>7</sup> It might therefore be said that the relations that hold between minds and bodies in Spinoza's philosophy are miraculous. Against this, Spinoza's doctrine of conatus might equally be taken as awarding a degree of natural causal efficacy to finite particulars too. This would bring Spinoza's account closer to that which I shall later attribute to Leibniz, with the difference, however, that Spinoza's conatus occurs in finite particulars that are mere modes of God's attributes rather than substances in their own right.

As Leibniz interpreted Descartes' interactionism, the theory conforms to option (iv): God arbitrarily determines the correlation between minds' sensations and bodies' motions and brings it about miraculously since Descartes held that God imparts and maintains motion in bodies (*Principles of Philosophy*: AT IXb.84, CSM I.240). On this view, Descartes may be read as a proto-occasionalist.<sup>8</sup> But if, as some commentators attest,<sup>9</sup> Descartes actually attributed force to bodies, his theory more closely instances option (iii): the mind-body relation is still the result of arbitrary divine decree and humanly unintelligible, even though the forces in minds and bodies enable them to act directly—and naturally—on each other.

Leibniz's claim is that pre-established psycho-physical correlations are intelligible to God and to humans, in some sense essential or necessary, and instigated by God acting through the natural, non-miraculous, inherent forces of minds and derivative forces of bodies. Leibniz conceived the link between the intelligibility of mind-body correlations and their arising by natural final and efficient causes as more than accidental. But, we may ask, does God's acting by natural, non-miraculous ways actually ensure that mind-body relations are essential or necessary and in principle humanly intelligible? Or, as in option (iii), might God act by natural means to bring about psycho-physical relations that are nevertheless thoroughly arbitrary and non-necessary? Conversely, does God's acting by miraculous means



inevitably lead to mind-body relations that are arbitrary and unintelligible to humans? Or, as in option (ii), might God act by miraculous means to effect psycho-physical relations that are essential and intelligible? Unless Leibniz can establish an inviolable link between intelligibility and God's acting by natural means, both options (ii) and (iii) remain possible. In what follows, discussion of the non-arbitrary character of the relation will allow us, in Section III, to highlight a certain convergence in the views of Spinoza and Leibniz, which will in turn illuminate how Leibniz rules out the Cartesian option (iii). We shall also note, though without further discussion, how Leibniz also counteracts the Spinozan option (ii), proposing instead a plurality of natural causes over a single supernatural cause.

While corresponding with Masham, Leibniz was also engaged in a detailed critique of Locke's *Essay Concerning Human Understanding*. In this, his *New Essays on Human Understanding* (NE), he touched on matters relating to the non-arbitrary character of the harmony of mind and body:

It must not be thought that ideas such as those of colour and pain are arbitrary and that between them and their causes there is no relation or natural connection: it is not God's way to act in such an unruly and unreasoned fashion. I would say, rather, that there is a resemblance of a kind—not a perfect one which holds all the way through, but a resemblance in which one thing expresses another through some orderly relationship between them. (NE II. viii: A VI vi, 131, RB 131)<sup>10</sup>

Colour and pain are related to their physical causes by means of a 'natural connection' that consists in a kind of resemblance and mutual expression based on an 'orderly relationship'. In his *Metaphysical Consequences of the Principle of Reason*, composed sometime during the period between 1695 and 1712, Leibniz explained his notion of 'expression' in terms of a 'constant relational law' that permits an isomorphic mapping of 'particulars' in the things expressing each other:

It is sufficient for the expression of one thing in another that there should be a certain constant relational law, by which particulars in the one can be referred to corresponding particulars in the other. (§11: C 15, MP 176–7)

In the case of mind and body, the expressive relation depends upon the soul having minute (*petites*)—or what Leibniz sometimes calls 'insensible'—perceptions of which it is unaware:

The insensible parts of our sensible perceptions also bring about a relation between those perceptions of color, heat, and other sensible qualities, and the motions in bodies that correspond to them. (NE Preface: A VI vi.55, RB 55, AG 296)



The role of these minute perceptions is not causal, but operates by means of an isomorphic mapping of the soul's perceptions to the movements of the body and its parts. Although perceptions and motions are radically different in kind, there are discernible affinities. For instance, Leibniz associated the 'continuing thrust of a heavy body gaining impetus as it falls' with the 'continuation and accumulation' of what he calls 'semi-pleasures' (that is, minute pleasurable perceptions) in the soul which, when amassed in sufficient numbers, 'eventually becomes a whole, genuine pleasure' (NE II. xx: A VI vi, 164, RB 164). Our pleasurable sensations are composed of little pleasurable perceptions in a way analogous to the acceleration or impetus of a moving body being composed of smaller impetuses or to the composition of an extended body as an aggregate of smaller extended parts.

In the *New Essays*, Leibniz also offered an analysis of colours and warmth in terms of their causes, namely, the motions of tiny particles that comprise coloured or warm objects (NE II.viii.13). Thinking it 'thoroughly reasonable that the effect should correspond to the cause' (NE II.viii.15: A VI vi, 131, RB 131)—our perceptions are so confused that it is impossible to rule out the possibility of such a correspondence (ibid.)—Leibniz speculated as to how the correspondence might work. The basic idea seems to be this: because our confused sense perceptions of physical objects are composed of an infinite multiplicity of *petites* or insensible perceptions, when we perceive a physical object, we not only perceive it confusedly as one whole thing, but we also perceive insensibly all of its infinitely many parts. Just as an extended physical object comprises a multiplicity of tiny extended particles assembled together, so too our confused perception of the object as coloured or warm comprises a multiplicity of tiny (or confused) perceptions assembled together. In theory, then, there could be a one-one correspondence between the constituents of the perception of the object and the constituents of the object itself. For example, although the confused sensation of pain 'does not resemble the movement of a pin . . . it might thoroughly resemble the motions which the pin causes in our body, and might represent them in the soul' (NE II.viii.15: A VI vi, 132, RB 132). In effect, the minute perceptions that coalesce to form the confused sensation of pain match the minute motions caused by the pin in the particles of the body.

Such a model, Leibniz suggested, grounds our tendency to locate the pain in our own bodies—for this is where the motion of the particles is found—rather than in the pin. The situation is rather different in the case of bodies' secondary qualities. These, we tend to attribute to bodies rather than to souls. For instance, we commonly locate light in the fire because the rapidly moving particles that our souls represent as the perception of light occur primarily in the fire:

That is why we say that the pain is in our body and not in the pin, although we say that the light is in the fire; because there are motions in



the fire which the senses cannot clearly detect individually, but which form a confusion—a running together—which is brought within reach of the senses and is represented to us by the idea of light. (NE II.viii.13: A VI vi, 131, RB 131)

Leibniz's position comes into sharper focus when contrasted with Malebranche's. Malebranche agrees that particular sensations are consistently correlated with particular bodily states. For instance, in the case of colour perception, sensations of the primary (or 'simple') colours are correlated with low-frequency vibrations in the eye. Sensations of the secondary (or 'compound') colours are correlated with higher frequency vibrations on account of their being composed of the frequencies of the vibrations of the relevant primary colours. Thus the sensation of green is correlated with frequencies that combine the frequencies associated with blue and yellow. White being the most compounded of all the colours (see *Search after Truth*, elucidation 16: OC III.258, LO 689), white sensations are matched to the highest frequency vibrations.

However, unlike Leibniz, Malebranche does not think that our sensations are themselves composed of minute perceptions of each of these frequencies, nor of minute perceptions of each of the individually vibrating particles. According to the Oratorian,

It is certain that the soul does not cause in itself all the perceptions it has of the objects surrounding it as soon as it opens its eyes and looks around in the middle of the countryside. For besides the fact that it sees them without willing to do so, it *has not the least knowledge of the construction of its eyes and brain, or of anything happening in them.* (*Search after Truth*: elucidation 17: OC III.326, LO 733, my emphasis)

Malebranche here denies that the soul has any knowledge of the finer parts of its body and infers from this that the soul does not produce its own sensations.

The situation is quite different in Leibniz's hypothesis of pre-established harmony. He rejected both the premise and the conclusion of Malebranche's argument.<sup>11</sup> He *did* think we have, if not knowledge, then at least minute perceptions of the organised parts of our bodies, through which we also indirectly perceive external bodies and their parts (to Arnauld, 30 April 1687: GP II.98, AG 87). The Leibnizian soul always perceives, albeit for the most part insensibly, everything that is happening in its body. Every vibration of each and every part of the body is meticulously mapped to a corresponding minute, insensible perception in the soul.

Leibniz maintained further that the soul is the cause of its own perceptions, including its perceptions of physical objects in the external world. The soul produces its own sensations 'in conformity with what happens in matter' (NE IV.iii.6: A VI vi.381; RB 381). It draws its perceptions out



from itself by means of its own appetitive force and in accordance with the law that generates its particular series of perceptions, all the time mirroring the changing states of bodies in minute detail:

[the] nature that pertains to the soul is representative of the universe in a very exact manner (though more or less distinctly), the series of representations produced by the soul will correspond naturally to the series of changes in the universe itself . . .

and conversely, the organisation of the body reflects the souls' perceptions, for it too

. . . in turn, has also been accommodated to the soul for the situations in which the soul is thought to act externally. (*New System*, GP IV.485, AG 144)

The representative nature of the soul is such that it must represent every last detail in the body. A substance is 'perceptive and representative of the whole universe, according to its point or view and according to the impressions (or rather the relations) its body receives mediately or immediately from all others' (to Arnauld, 30 April 1687: GP II.98, AG 87) (see also *Monadology*, §60; GP VI.617). If the movement of my arm requires the movement of every part of my arm, then my perception of my arm moving will also involve a perception of each part of my arm as well. Taken together, these minute perceptions constitute the soul's conscious sensation of the body and its qualities. The confused perception of my arm moving is composed of the minute perceptions of the movement of the parts. The correspondence between perceptions and motions envisaged by Leibniz is thus far from superficial and arbitrary. Its intelligibility is grounded in the representative nature of the soul and the thoroughgoing isomorphic mapping of its perceptions to the motions of its body, a mapping that extends deep into the recesses of the soul and to the insensible parts of our bodies. This, Leibniz told Masham, is not 'merely possible in itself, but is also most in conformity with God's wisdom and with the order of things' (Leibniz to Masham, 30 June 1704: GP III.354, WF 212).

There is thus a rationale and implicit intelligibility to the relation between our sensations and the motions of our bodies. There is also a rational intelligible relation between our sensations and the inner structures of external bodies. An example from Leibniz's discussion on the Catholic doctrine of the Eucharist in his correspondence with Des Bosses helps illustrate his view. There, Leibniz postulated that God might miraculously break the intelligible link between the soul's sensations and the inner structures of bodies so as to make Christ's body appear to the soul as bread. Substituting whiteness and blackness for the appearances of the bread and the body of Christ, he imagined how God might give the soul the sensations



corresponding to a white object while the object itself had the structure of a black object. The soul would have the sensation of white without also having insensible perceptions of the ‘foam or little mountains’ that would normally accompany the sensation. In their place, the soul would have minute perceptions of ‘little valleys’, the usual cause of the appearances of black things:

If God wanted to substitute a black thing for a white thing, preserving the accidents of the white thing, he would bring it about that all perceivers (for in the mutual agreement of perceivers consists the truth of the phenomena) retain the observed perception of the white thing and its effect, that is, the perception of what results from that which constitutes the white thing. But they wouldn’t have the unobserved perception of the foam or the little mountains (that is, the textures producing a white thing); rather, they would have the unobserved perception of valleys, that is the unobserved perception of textures producing a black thing. (to Des Bosses, 29 May 1716: GP II.521, AG 206)

The substitution would, however, be miraculous. The normal situation is that the sensation of white arises in the soul from its minute perceptions of the foamy balls or little mountains that comprise the inner structure of the white object. In similar fashion, the norm is for the sensation of black to arise in the soul from its minute perceptions of the little valleys that comprise the inner structures of black objects.

But could God have made it the norm that the perception of white is composed of insensible perceptions of little valleys instead of mountains? Is the actual correlation just decreed arbitrarily? Quantum mechanics has since cast doubt on the notion that the visible and the sub-atomic realms operate on similar principles, but to the early modern mind, the correlation of the perception of bright white with light reflected from the tops of the ‘little mountains’ at the atomic level and the correlation of perceived dark blackness with light being trapped in the ‘valleys’ would certainly have appeared more ‘intelligible’ than the alternative.<sup>12</sup>

However, Leibniz claimed not only that the relation is intelligible, but also that it is natural. By this, he means souls and bodies act by their own forces, without direct miraculous intervention from God. Without minute perceptions, souls would not have confused and distinct perceptions—just as bodies would not have the gross qualities they have if they did not have their smaller parts as well. They need these because souls and bodies are the causes of their own changes. In order to be so—in order for the changes in motion and the changes in perceptions to arise from bodies and souls themselves, without direct and miraculous intervention by God—the minute perceptions in souls and the tiny parts of bodies are essential as the building blocks from which are formed souls’ sensations and bodies large enough to be visible to us. Minute perceptions and the intricate organisation



of bodies are the means by which God allows souls and bodies to act by their own forces, not by His. In this way, God operates fittingly by natural means, not by miraculous means, to bring about a rationally ordered, intelligible correspondence between mental and physical states.

God always acts in the way which is most in conformity with his perfections . . . And it is obvious that nothing is more excellent or better arranged than this prior agreement which God has established in natural things. (Leibniz to Masham, 30 June 1704: GP III.354, WF 212)

While the intelligibility and naturalness of pre-established mind-body relations are sufficient to answer Masham's criticism, in the following and final section, I suggest that Leibniz conceived a strong relation between (non-interactionist) naturalness and intelligibility. The mind-body relation is intelligible because it is necessary or essential and it is so precisely because it arises necessarily or essentially from the natural, but non-interactionist, means God chose to effect the relation.

### III

Leibniz's monads are active and passive primitive forces. When the monad is created, its law of the series—the individual law that contains details of all the changes the individual substance will undergo in the course of its life—unfolds as a series of perceptions, brought about by individual appetitions. These perceptions and appetitions are described as modifications or temporary fleeting states of the primitive force, where the appetite is the activity of the force that moves the monad from one perception to the next. At the same time, the monad's primitive force is also modified as the temporary fleeting derivative forces that exist as its organic body. The presence of derivative forces in bodies is indicated by the body's motion, in the case of derivative active force, and by the body's resistance, in the case of derivative passive force. The modification of monadic primitive force as physical derivative forces is an external modification and occurs only on condition that God has actually created the other substances that comprise the dominant monad's organic body.<sup>13</sup> The main point to note here is that Leibniz's monads are essences that are simultaneously modified both as a series of mental states (perceptions and appetitions) and as a series of physical states (derivative forces manifested as the body's motion and resistance).

On this model, Leibniz's finite substances bear a striking resemblance to Spinoza's infinite Substance, God. Both thinkers conceived substance (in the case of Spinoza) or substances (in the case of Leibniz) as entities whose essences give rise concurrently to two different sets of modes: the one, mental ideas or perceptions brought about by appetitions; the other,



physical bodies or derivative forces. For Spinoza, God is both a thinking thing and an extended thing. God's essence is simultaneously modified as ideas, insofar as God is a thinking thing, and as simple and composite extended bodies, insofar as God is an extended thing.<sup>14</sup> Being modifications of the same essence, it is assured that the order of ideas corresponds to the order of causes: 'The order and connection of ideas is the same as the order and connection of things' (*Ethics* II, Proposition 7: Curley 451).

Leibniz's pluralist metaphysics advances a pluralism of finite substances, whereas Spinoza's monism advances a single infinite Substance and regards finite individuals as mere collections of modes of this one Substance, God. So too, Leibnizian bodies are aggregates of substances rather than collections of modes deriving from God's essence. Nevertheless, Leibniz's finite individual substances resemble Spinoza's infinite Substance in this respect: for Leibniz, each created individual has an essence (or primitive force) that is simultaneously modified as a sequence of appetitions and perceptions and as the repetition of derivative forces that comprise its organic body and ensure that it exists as a living corporeal substance that is both a perceiving being and a physical thing. For Leibniz, as for Spinoza, the cause of the sequence of perceptions and the cause of the movements of the body are one and the same—in Leibniz's case, the essence or primitive force of the individual substance itself. And because each set of modifications, perceptions, and derivative forces arises from the same individual essence, the correlation between the two series is assured. Moreover, because the derivative forces of bodies are manifested as the motions and resistances of bodies, the correlation between perceptions and motions is also assured. Hence, just as for Spinoza, the order of ideas corresponds exactly to the natural order of bodies because both arise from the same infinite Substance, so too, for Leibniz, the order and connection of a finite substance's perceptions exactly match the order and connection of its body and all its parts because both arise as modifications of the essence or nature of the substance's essence or primitive force.

For Spinoza, the correlation of mind and body is necessary for both mind and body derive from one and the same essence. Given the similarity to Leibniz's account of the dual modification of finite substances, it is not unreasonable to postulate that for Leibniz, even though the actual existence of particular souls and bodies is only hypothetically necessary—that is, dependent on God's choice of the best possible world—it is possible that the correspondence relations themselves are absolutely necessary, such that, given God's decision to create these particular souls and a determination to create bodies that correspond to their perceptions, which bodies are actually created may then be dictated by an absolute necessity. In other words, it may be that for Leibniz, the correspondence between sensations and physical states holds with a necessity equally as strong as that proposed by Spinoza. And even if the necessity is not quite so strong and God could have chosen to create a less-than-perfectly-ordered world in which



different correlations obtained, Leibniz has shown that God's acting by natural means does not have to be understood along the interactionist lines that led Descartes to propose that the mind-body relations are established by God by arbitrary decree (option (iii)). Psycho-physical relations can both arise from natural causes and be thoroughly essential and intelligible.

For both Leibniz and Spinoza, then, the mind-body correspondence is essential, for it arises from the essences of substances. It is also in principle intelligible—in both cases, a reason can be given why the correlations are as they are. But the Spinozan option (ii) postulates that the correlation is effected directly by God. It is, to all intents and purposes, miraculous. However, by insisting that there is a plurality of substances, each acting by their own natural forces, Leibniz has shown that it is possible that the immediate cause of bodily actions and perceptions is not God, but the individual substance itself—a position advantageous to morality in making room for individual freedom of will and responsibility. Moreover, since it is the same natural cause of both sequences, the intelligible correspondence between the two sequences is also natural (option (i)). Far from being arbitrary, miraculous, and incomprehensible to finite minds as Descartes (option iii) and Malebranche (option iv) supposed, or resulting solely from the essence of God as Spinoza maintained, the mind-body relation in Leibniz's system is grounded in reason, natural forces of individuals, and intelligible even to humans. Moreover, answering Masham, the organisation of the body is essential to matter in order for this to be so.

## NOTES

1. The following abbreviations are used in this chapter: A: G. W. Leibniz, *Sämtliche Schriften und Briefe*, 7 vols., edited by Deutsche Akademie der Wissenschaften (Darmstadt/Leipzig/Berlin: Akademie Verlag, 1923–); AG: G. W. Leibniz, *Philosophical Essays*, edited and translated by Roger Ariew and Daniel Garber (Indianapolis and Cambridge: Hackett, 1989); AT: R. Descartes, *Oeuvres de Descartes*, 12 vols., edited by C. Adam and P. Tannery (Paris: Vrin, 1964–76); C: G. W. Leibniz, *Opusculs et Fragments Inédits de Leibniz*, edited and translated by L. Couturat (Hildesheim: Olms, 1903, reprinted 1966); CSM: R. Descartes, *The Philosophical Writings of Descartes*, 2 vols., edited and translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge: Cambridge University Press, 1985); D: G. W. Leibniz, *Opera Omnia*, 6 vols., edited by L. L. Dutens (Hildesheim: Olms, 1768, reprinted 1989); GP: G. W. Leibniz, *Die philosophischen Schriften*, 7 vols., edited by C. I. Gerhardt (Berlin: Weidmannsche Buchhandlung, 1875–90); JS: N. Malebranche, *Dialogues on Metaphysics and on Religion*, edited by Nicholas Jolley, translated by David Scott (Cambridge: Cambridge University Press, 1997); LO: N. Malebranche, *The Search After Truth*, edited and translated by Thomas M. Lennon and Paul J. Olscamp (Cambridge: Cambridge University Press, 1997); MP: G. W. Leibniz, *Philosophical Writings*, edited and translated by M. Morris and G. H. R. Parkinson (London: Dent, 1973); OC: N. Malebranche, *Oeuvres complètes de Malebranche*, edited by André Robinet, 20 vols. (Paris: J. Vrin, 1958–76);



RB: G. W. Leibniz, *New Essays on Human Understanding*, translated by Peter Remnant and Jonathan Bennett (Cambridge: Cambridge University Press, 1982, reprinted 1985); WF: G. W. Leibniz, *Leibniz's "New System" and Associated Texts*, edited and translated by Roger Woolhouse and Richard Francks (Oxford: Clarendon Press, 1997).

Previous versions of this chapter were presented at conferences in Neuchâtel (2008), London (2008), and Aberdeen (2010). I am most grateful for helpful comments from participants at each.

2. 'I have always observed in astronomy and in physics that the most intelligible theories prove true in the end; for example that of the movement of the earth to save the appearances of the stars, and that of the weight of the air to explain air pumps and other attractions which had formerly been attributed to the abhorrence of a vacuum' (Leibniz to Masham, 30 June 1704: GP III.353, WF 211).
3. See also *New System* (GP IV.486, AG 145) and postscript to a letter to Basnage de Beauval, 3/13 January 1696 (GP IV.499, WF 63).
4. God's immediate action on human minds can be regarded as communication between heterogeneous substances only if we accept the Cartesian view that the 'term "substance" does not apply univocally' to God and to created things (*Principles of Philosophy*, part 1, §51, CSM I, 210). Stuart Brown notes that prior to 1686, Leibniz seems to have believed that minds can effect changes in bodies. See S. Brown, "Malebranche's Occasionalism and Leibniz's Pre-established Harmony: An 'Easy Crossing' or an 'Unbridgeable Gap'?", in *Nicolas Malebranche: His Philosophical Critics and Successors*, edited by S. Brown (Assen, Netherlands: Van Gorcum, 1991), 81–93, at 82.
5. There are moral reasons why God gives souls sensations. As well as enabling the identification of distinct objects, sensations of light and colour also highlight relations among objects relevant to 'the convenience and preservation of life' (*Dialogues on Metaphysics and on Religion*, Dialogue 5: OC XII.113; JS 74).
6. The body's derivative force is a modification of the primitive force of the body's animating soul if it is a corporeal substance or, in the case of inanimate bodies, by a collection of such modified primitive forces.
7. Even if we stress Spinoza's equation of God (*natura naturans*) with Nature, we cannot infer from this that the correspondence is 'natural' in the sense used here (*natura naturata*).
8. See, for instance, Steven Nadler, "Descartes and Occasional Causation", *British Journal for the History of Philosophy*, 2 (1994) No. 1: 35–54.
9. For example, Dan Garber, *Descartes' Metaphysical Physics* (Chicago & London: University of Chicago Press, 1992), 294f.
10. Presumably Leibniz had Descartes and Malebranche as well as Locke in mind when he remarked: 'Some contemporaries believe that our ideas of sensible qualities differ entirely from motions and from what occurs in the objects, and are something primary and inexplicable and even arbitrary; as though God had made the soul sense whatever he had a whim that it should sense, rather than whatever happens in the body—which is nowhere near the right analysis of our ideas' (NE II. xx: A VI vi, 164, RB 164). Besides, it would be unworthy of God 'who does nothing without harmony and reason' not to take note of 'any essential relation between perceptions and their objects' (NE Preface: A VI vi, 55, RB 55, AG 296).
11. The suppressed premise of Malebranche's argument, namely, that 'something can only affect something if it knows it', is replaced in Leibniz's system by the claim that something can only be said to act on something else if it perceives the other more distinctly than the other perceives it.



12. A similar thought underlies Leibniz's example of the sound of the sea as composed of the sounds of each breaking wave (*Discourse on Metaphysics*, sect. 33: GP IV.459, AG 65).
13. I have argued this elsewhere. See P. Phemister, *Leibniz and the Natural world: Activity, Passivity and Corporeal Substances in Leibniz's Philosophy* (Dordrecht, Netherlands: Springer, 2005), 194–202.
14. Arising directly from God's essence, ideas and bodies may be said to arise from God's miraculous action, as in option (ii), although Spinoza also maintains that modes are, in the order of nature as *natura naturata*, caused by each other.



## 6 Hobbes's Redefinition of the Commonwealth

*Timothy Stanton*

Hobbes's definition of the commonwealth has two claims to appear in this volume: as involving positions about the causation of commonwealths, and as contributing to the causal conditions which to his mind sufficed for their existence. The first claim arises because Hobbes insisted that definitions, which on his account formed the basis of scientific demonstrations, needed to be such as to express the causal generation of the things they defined: 'where there is place for demonstration [he wrote in one place] if the first principles, that is to say, the definitions contain not the generation of the subject, there can be nothing demonstrated as it ought to be' (SL: 184).<sup>1</sup> The second claim arises because with him words had causal power. That is to say, they affected the world directly. These theses flowed from Hobbes's wider philosophical positions, amongst which his metaphysical positions were especially important.

'Metaphysics' is used here in an Aristotelian sense, meaning first philosophy. First philosophy is that science, separate from the other sciences, concerned with the establishment of those sciences as separate and distinct, on the basis of their first principles. The world and its wife knows that Hobbes found much to dislike in Aristotle, which he delighted in mocking unmercifully (e.g., L: 458–74);<sup>2</sup> but he was apparently impressed enough by Aristotle's *Metaphysics* to elaborate his own first philosophy according to the pattern of explanation developed there.

On that pattern the first principles of each distinct science are definitions. Aristotle suggested that his many predecessors had failed to establish these sciences upon their proper foundations because of the inadequacies of the definitions they had propounded and the methods they had used in making them, and he concluded that, for want of method, those definitions had failed to account for the generation of their subjects.<sup>3</sup> In his turn, Hobbes suggested that there was 'a certain *Philosophia prima*, on which all other Philosophy ought to depend' and that it consisted 'in right limiting of the significations of such Appellations, or Names, as are of all others the most Universall . . . [which] are commonly called Definitions' (L: 463).<sup>4</sup>



As he put it elsewhere, 'the making of definitions, in whatsoever science they are to be used, is that which we call *philosophia prima*' (SL: 222). He also suggested that 'want of Method' had capsized his predecessors' overwhelmingly dismal efforts at constructing sciences proper (L: 34), not least because *their* definitions had failed to provide for the generation of their subjects.<sup>5</sup> The grand tripartite scheme of knowledge which Hobbes envisaged was constructed on this ground: the three 'Elements of Philosophy'—*De corpore*, *De homine*, *De cive*—comprising, respectively, the distinct sciences of body, man, and citizen, with each proceeding by the method of reckoning—adding and subtracting—the consequences of names on the basis of first principles—definitions—established in terms of the first philosophy presupposed by all.

The starting point of each science so conceived was, then, a small collection of definitions. For unless someone knew what the names he used stood for, he could not possibly reckon their consequences rightly. He would find himself, in Hobbes's memorable phrase, 'entangled with words, as a bird in lime-twiggs; the more he struggles, the more belimed'. Thus it was imperative to settle the signification of words before attempting to reckon them up. By the same measure, any man that aspired to scientific knowledge would need 'to examine the Definitions of former Authors; and either . . . correct them if they are negligently set down; or . . . make them himself. For the errors of Definition multiply themselves, according as the reckoning proceeds; and lead men into absurdities, which at last they see, but cannot avoyd, without reckoning anew from the beginning; in which lies the foundation of their errors' (L: 28). This pattern of negligence and error seemed to Hobbes (as it had seemed to Aristotle) more or less ubiquitous, but under the pressure of events his attention centred more and more upon one particular instance of it, the definition of the 'commonwealth'. His political theory presented a redefinition of the commonwealth, a definition he duly positioned at the apex of the civil science—or 'civil philosophy'—he claimed to have originated.<sup>6</sup>

This essay argues that a consideration of the significance of this definition is in order, because it has philosophical bearings which affect Hobbes's account of politics in ways that are fundamental but seldom addressed by commentators. If Hobbes resembled Aristotle in his view of first philosophy, he otherwise thought in very different philosophical terms. These terms affected the significance of the definitions he proposed and conditioned his attitude to those he wished to supplant and to silence: for they indicated that the right settling of definitions was not merely a matter of speculative importance, but rather a matter of ineluctably practical importance. To see why, it is necessary to begin with those 'former Authors' whose definitions of the commonwealth Hobbes disliked and looked to replace with his own, before turning in the later sections of the essay to Hobbes's conception of causality and its consequences in his political thinking.



## I

A suitable place to begin is Sir Thomas Smith's *De Republica Anglorum*, in which a commonwealth is defined, famously, as 'a societie or common doing of a multitude of free men, collected together, and united by a common accord & couenants among theselues, for the conseruation of themselves as wel in peace as in warre'.<sup>7</sup> Smith's definition established a categorical distinction between commonwealths, which were voluntary associations of free individuals, and other forms of association, whether that between mere hosts of men who were not united by common accord or that between masters and slaves who, because they were not free and had no communion with their masters, did not constitute a commonwealth with them.

The problem with this definition, as Hobbes saw it, was that it involved a crucial elision. To be specific: Smith's definition stipulated that common accord could produce unity: that is, that the former was a sufficient condition of the latter. Hobbes denied that this was so. He sought to further distinguish between what Smith had called 'common accord' and he himself called variously 'consent', 'accord', and 'concord' (in Latin, *consensio* or *concordia*) and what he called 'union' (in Latin, *unio*). Hobbes argued that union and covenant were correlated but that union and common accord were not, and so that the second was insufficient to produce the first (compare EL: 72, DC: 71–72, L: 120).<sup>8</sup>

This argument appeared more or less fully formed in the earliest presentation of his political theory, *Elements of Law* (EL: 84), but it was elaborated more systematically in *De cive*. There Hobbes observed that an accord between several parties consisted 'only in their all directing their actions to the same end and to a *common good*'. But, he continued, 'something more is needed . . . to prevent an accord on peace and mutual assistance for a *common good* from collapsing in discord when a *private good* subsequently comes into conflict with the *common good*'. The 'something more' was 'an element of fear'—always, with Hobbes, the passion to be reckoned upon—and this element was introduced by the creation or common acknowledgement of 'a *single will* among all of them in matters essential to peace and defence. This can only happen [Hobbes went on] if each man subjects his *will* to the will of a *single* other, to the *will*, that is, of one *Man* or one *Assembly*, in such a way that whatever that one *wills* . . . may be taken as the *will* of all and each' (DC: 70–71).<sup>9</sup>

This general subjection Hobbes called 'UNION' and the union so made a '*commonwealth* (*civitas*) or *civil society* and also a *civil person*; for since there is *one will* of all of them, it is to be taken as *one person* . . . A COMMONWEALTH, then (to define it) is *one person*, whose *will*, by the agreement of several men, is to be taken as the *will* of them all; so that he may make use of their strength and resources for the common peace and defence.' (DC: 73) In *Leviathan*, Hobbes used the language of representation to explain the same operation, adding that it was in this person



that 'consisteth the Essence of the Commonwealth' and that whoever bore this person was 'SOVERAIGNE, and said to have Sovereigne Power; and everyone besides, his SUBJECT' (L: 121). In short, a commonwealth was the union of all in one person, the sovereign. A commonwealth without sovereignty, by contrast, was not a commonwealth at all.

If Smith's definition had failed to register this fact, it was not alone. The classical notion of a mixed constitution, recently revived by supporters of the king who wished to resist pretensions to parliamentary infallibility without simply inverting them, embodied the same failure.<sup>10</sup> For the essence of a mixed constitution was that its different elements were, in Hobbes's terminology, in accord, as opposed to united.

That this was a failure of *definition* Hobbes was explicit: 'The error concerning mixed government', he wrote in *Elements of Law*, 'hath proceeded from want of understanding of what is meant by . . . *body politic* [*sc.*, commonwealth], and how it signifieth not the concord, but the union of many men' (EL: 167). A commonwealth and a mixed constitution were therefore antithetical to one another, for without the union of all in a single person through their submission to that person, a group of men was simply a group of men, a multitude of individual wills; and a multitude, whether on Hobbes's definition or on Smith's, was not a commonwealth.<sup>11</sup> The fact that these men directed their actions 'to the same end and to a common good' was, we might think, a creditable achievement—though Hobbes attributed the same capacity to ants and bees—but it was insufficient to constitute a civil association (L: 86).

At this point it might be objected that the 'something more' Hobbes introduces into Smith's definition of the commonwealth surreptitiously destroys what is most distinctive about it, namely the freedom it attributes to the men who compose the commonwealth. Again in *Elements of Law*, Hobbes had argued that 'the subjection of them who institute a commonwealth amongst themselves is no less absolute, than the subjection of servants' (EL: 132). This was not without polemical edge: for the argument that in their civil associations free men were positioned no differently than servants or slaves was surely calculated to offend those opponents of the king for whom the categories of free man and slave were mutually exclusive; and it would have been the more provoking still for being repeated *mutatis mutandis* in a book called *De cive*, where despotic subjection was discussed under the heading of citizenship (compare EL: 126–27, DC: 111–12).<sup>12</sup> Quentin Skinner is only the most distinguished modern commentator to conclude that by this and like arguments Hobbes was endeavouring to effect the reduction of the 'commonwealth' to the simple acknowledgement of absolute subjection.<sup>13</sup>

A similar conclusion was earlier reached by one of Hobbes's contemporaries, the veteran royalist pamphleteer Sir Robert Filmer. This was what made Hobbes's position appealing in his eyes. Writing in 1652, he reported that it had been '[w]ith no small content [that] I read Mr. Hobbes book *De*



cive, and his *Leviathan*, about the rights of sovereignty, which no man, that I know, hath so amply and judiciously handled'.<sup>14</sup> But Filmer followed up his praise first of all with a qualification and then with an expression of disagreement.

The qualification he entered was that, given that Hobbes had correctly identified the rights of sovereignty, he had attached the wrong name to the civil association constituted around them. According to Filmer, the name 'commonwealth' connoted 'a popular government, wherein wealth and all things shall be common'; whereas the 'true' name for the mode of association Hobbes was depicting was rather 'weal public, or commonweal', the name preferred in Richard Knolles's influential translation of Bodin's *De Republica*.<sup>15</sup> Plainly, Hobbes's choice of terms was careful—why, then, did Filmer think it mistaken? It is probable that Filmer had ideological reasons in 1652 for wishing to distance Hobbes's arguments from the idea of a 'commonwealth'.<sup>16</sup> But there were also philosophical reasons, which followed from Filmer's reliance on certain Aristotelian assumptions: because he thought about some matters in the same philosophical terms as Aristotle,<sup>17</sup> Filmer differed from Hobbes about the character and definition of the commonwealth.

## II

Filmer relied on Aristotelian views about causality and nature as postulates in elucidating his claims about the political authority of sovereigns. These views hardened his opinions about how sovereign power should and should not be defined. They led Filmer to reject Hobbes's definition of the commonwealth, because he was unable to countenance—indeed, he found obnoxious—the explanation of how sovereign power came about which that definition incorporated. Filmer thought that it was impossible for free men to submit themselves to a single will, because men were not free to begin with. Rather, they were born into subjection, in the sense that they were naturally subject to the authority of another. Filmer traced this subjection back to the first man, Adam, who enjoyed authority over his family and all their posterity. He explained his authority with reference to God and to Aristotle.

By Filmer's account Adam enjoyed authority over his family in part because he had generated them. In his words, God, having created Adam, took 'a piece of him' and 'made the woman', and then 'by generation from them two as parts of them all mankind' were propagated.<sup>18</sup> But why should the generation of the rest of mankind out of Adam be thought to give him authority over them? In answering this question, God joined hands with Aristotle.

Aristotle's *Generation of Animals* stated that in procreation the male was active and the female merely the passive receptacle of his activity.<sup>19</sup>



This distinction of roles suggests that in some sense the male is more important than the female in procreation. As Filmer himself wrote, it was the man that was 'the nobler and principal agent in generation'—and this was what gave him 'sovereignty . . . over the woman'.<sup>20</sup> How exactly did it do this? Here we encounter Aristotle's four causes.

Nowadays we typically understand causality as Hume understood it: as a connection between two things in which the first is necessary or sufficient for the second to occur, as when one billiard ball strikes another and causes it to move.<sup>21</sup> On this understanding, a cause is an event antecedent to and necessarily connected with another event called the effect. Aristotle took a broader view of causality, which embraced four distinct components in a single understanding. In his view, a cause could be any one of the components which explained why a certain thing was or came to pass.<sup>22</sup>

Aristotle identified, first of all, the matter out of which something comes, as a figurine from clay. This was the material cause. The material cause he distinguished from the formal cause. The formal cause of a thing was not the material from which it was made but rather its essence—we find Hobbes writing of 'the ESSENCE, which some call the *formal cause*' (DCO: 131)—in other words, the abstract pattern or form to which the material would correspond in its finished state. It was the formal cause that gave a thing the character that differentiated it from the matter out of which it came; through it, a mass of matter became something else and more than that. Matter and form, then, together constituted an entity. What brought that entity into existence? This role was attributed to the efficient cause, the source of the movement or impulse by which that entity originated. Lastly, there was the final cause, which Aristotle understood as the end or purpose for the sake of which the originator of an entity brought it into existence.

On this view, in order to understand an entity it is necessary to consider all of its component causes, as all four may enter into the explanation of it. For example, a figurine requires not only the clay out of which it is made as the material cause and the efficient cause of the modeller, but also the formal cause of the principles of representation by which he is working and the final cause of the purpose for which he has created the figurine.

How does this relate to generation? It suggests, first of all, that two or more causes would be needed to produce any effect whatever. What of their content? On an Aristotelian understanding, the male's activity in procreation encompassed both efficient and formal causality, while the female's passivity manifested itself in material causality—the male 'acts and makes, while that which is made and receives the form' is made from material contributed by the female.<sup>23</sup> Hence the superior importance of the male to female in procreation: while the female provided the matter upon which the male acted as an efficient cause, the male was both efficient and formal cause, providing the initial impulse for their offspring and also communicating its essence. On these assumptions, it is easy to see why Filmer should have identified the male as the 'principal agent' in generation, because the



male plays two roles to the female's one.<sup>24</sup> The superior nobility of his role is also evident, for it is a more elevated activity to provide the formal cause that turns matter into something new than to supply the matter itself. In human procreation, the male's role was the more important, because the 'essence' of a human being was the soul;<sup>25</sup> and so his role involved supplying the constituents of the soul and in doing so bringing a new human being into existence. Thus Filmer spoke of this process as 'creation' and went on to say that it was this creative power that made men 'princes over their posterity'.<sup>26</sup>

The move from creative power to authority was completed via the scholastic maxim that the cause is greater than the caused, together with the idea that causal superiority implied authority when it was united with eminent intelligence and goodness. God was the great exemplar and embodiment of this combination, for His attributes—of omnipotence, omniscience, and goodness—combined to produce a power of creation (through the first of these attributes) and a claim by nature to direct others (through the second and third).<sup>27</sup> If some human beings conspicuously lack these attributes, it was not thought to be so with Adam. His claim to creative power was evident because he brought his children into existence, while his claim to direct others followed from the same nexus of ideas: for as a creator he was assumed to know intimately the thing that he had created, just as God understood His creation intimately. Consistently with this it was widely supposed that Adam understood not only the essence of human beings but also the essences of all the plants and animals of the Garden: so Alexander Ross could refer to his 'knowledge and wit, in giuing names to the creatures according to their natures', adding that he named them 'in token of his power ouer them'.<sup>28</sup>

Adam's authority, then, was natural by virtue of his procreative role and other characteristics.<sup>29</sup> But Filmer went further. He interpreted paternal authority over children as political authority, and thus presented Adam as lord paramount over his descendents, that is to say, the whole human race.<sup>30</sup> His dominion was not only political but proprietorial, since God had given to him 'not only the dominion over the woman and the children that should issue from them, but also over the whole earth to subdue it, and over all the creatures on it' (rendering thereby the hypothesis of natural freedom as impious as it was false).<sup>31</sup> After Adam's death, his dominion devolved to his male descendents because, as Filmer argued elsewhere, Adam stood for all mankind in a representative capacity, such that 'what was given unto Adam . . . was given in his person to his posterity'.<sup>32</sup> Or as Edward Stillingfleet put it, Adam was 'a publick Person by Representation, and not merely by Nature'. It followed that what he received for himself—in this case, dominion—he received not only for himself but also for all whom he represented. For himself he received sovereignty over the rest of the human race; and because he united in his person the roles of husband, father, and king, the authority he enjoyed in each role in his representative capacity was



bequeathed to every occupant of these roles.<sup>33</sup> In fact, the roles were to all intents and purposes interchangeable, since Filmer's combination of causal and juridical superiority produced a like authority in each.<sup>34</sup> The result was that in creating his family Adam had created the first political society. Hence Filmer's assertion that there was, properly speaking, 'no other sort of government, but only monarchy', because the essence of sovereignty lay in the dominion transmitted through Adam as husband, father, and king to all subsequent monarchs.<sup>35</sup> Hence also his preference for the name 'common-weal' to describe that political society, a term which denoted the rule by one fitted to direct others to their collective benefit. We can now see why Filmer would have wished to qualify Hobbes's terms, since the Hobbesian sovereign was said to possess the requisites needed to direct others and to exercise those requisites in the interest of those directed (see L: 121–7, 175–86). In any event, Filmer's view was that a political society was created through a particular transaction in which efficient, material, formal, and final causes each had their own parts to play, and that its 'true' definition would contain a proper account of its generation: accordingly he cited with approval Plato's definition of the commonweal as 'nothing else but a large family'.<sup>36</sup>

It is worth pausing to add that Smith's definition of the commonwealth presupposed a similar model, structurally speaking. His definition contained an efficient, material, final, and (at least notionally) a formal cause, insofar as it identified the agency by which a commonwealth was brought into existence, the material out of which it came, the end for which it was brought into existence, and the essence—the concord of free men—that distinguished it from the material from which it was composed. Restated in these terms, Hobbes's objection to Smith's definition may be said to be that it lacked an adequate formal cause, and hence that his commonwealth collapsed back into a multitude: for without a formal cause it was indistinguishable from the material from which it came.

The surpassing point for present purposes, however, is that with Filmer and Smith alike what we have is an explanation of how a particular society is generated via the communication of an essence. When we turn to Hobbes's definition of the commonwealth, it is tempting to conclude that a similar transaction is being recounted, in which men subject their wills to the will of a single other in order to generate or institute the sovereignty which is essential to the existence of the commonwealth that secures their common peace and defence. This temptation was irresistible to Filmer, who understood Hobbes's definition to be inadequate on precisely those grounds: the transaction which Hobbes identified, he complained, never could have occurred. What is more surprising, perhaps, is how irresistible it has been for modern commentators, who have discovered in Hobbes's political writings what they suppose is the attempt to explain the terms on which individuals transact with one another to construct a commonwealth out a state of nature.<sup>37</sup>



There is a difficulty with this supposition. The difficulty is that, on Hobbes's understanding, such a transaction is impossible. Whatever else it might involve, generation *cannot* involve the communication of an essence. It cannot do so because generation did not involve the same ideas for Hobbes as it did for Filmer. The reason for the difference is that Filmer's account employed the logic of Aristotle's understanding of causality. Hobbes's account did not. Hobbes's conception of causality took over Aristotle's causes but imparted to them a distinctive character of his own, which combined with other elements in his thinking to move his account beyond Filmer's, onto a different plane of explanation. This combination of elements helps us to understand some features of his argument which have perplexed Hobbes's readers from his day down to our own. The next section of this essay considers the first of these elements, Hobbes's conception of causality.

### III

What is a cause with Hobbes? A cause is a combination of properties in two things, one of which—the agent—acts and the other of which—the patient—is acted upon. These properties must either be necessary or sufficient for bringing about the effect that they produce; and insofar as they *are* necessary or sufficient for bringing that effect about, they form a single 'integral' cause of the effect (AW: 314). At first glance, this conception looks rather more like Hume's than it does Aristotle's, but a more searching glance suggests that, in his own way, Hobbes was responding to the conceptual agenda set by the four causes.<sup>38</sup>

On a second look we can see that Hobbes's conception involves at least two of Aristotle's four causes, what he had called the efficient cause and the material cause. For in Hobbes's words, 'all properties present in the agent, if taken collectively', could be termed the 'efficient cause' and those in the patient the 'material cause'. So far, this account sounds very much like Filmer's Aristotle.<sup>39</sup> Hobbes conceded the similarity himself when, having discussed these two causes, he added that it would be 'usual to point to two other kinds of cause: the formal and final' to complete the picture (AW: 314–15). But at this point, as at other points, Hobbes preferred to be unusual. How did he deal with the other two kinds of cause?

About final causes, Hobbes was emphatic: a final cause 'as far as a man can understand it, is exactly the same as an efficient case'. That is to say, it is a property of an agent that is hypothetically necessary for the effect it produces. The 'final cause', the end or purpose for which an agent acts, is, for Hobbes, merely the last of a series of antecedent causes which sets in motion a further series in which the desire of obtaining that end is the efficient cause of its movement towards the end it has in view—or as he put it, 'from something pleasant arises the thought of enjoying it; from the



thought arises the notion of the path to secure it; and from the notion of the path arises the progression towards the object desired'. In other words, final causes were simply efficient causes misnamed: they neither identified a distinctive type of cause nor belonged to a different order of causation (AW: 314–15).

If final causes reduced to efficient causes, what did Hobbes make of formal causes? The answer is, not much: at one point he stated that a formal cause was 'correctly speaking, not a cause' at all (AW: 314). What did he mean by this statement? With Aristotle, to recollect, the formal cause (the 'essence' or 'nature' of a thing) was what made that thing what it was and distinguished it from the material that composed it. Now there is a sense in which Hobbes agreed with this—often he appealed to essences or natures in his own explanations of what made things what they were. What he denied was that they possessed the character and causal status attributed to them on the Aristotelian view.

In elucidating his own view, Hobbes invited his readers to imagine a geometrical figure such as a 'square'. Because it is a square, all its internal angles are right angles. This being so, its being a square was said by those of Aristotle's view to be the cause of the 'right-angledness' of its internal angles. But what kind of cause could it be? As these same people acknowledged, it could not be an *efficient* cause, because its being a square was simultaneous with its having right angles—there was no prior action on the part of the square that changed its state from one of 'non-right-angledness' to 'right-angledness'. So they called it a 'formal' cause instead, because the 'right-angledness' that resulted was said to be a consequence of the 'form' of the square rather than being a consequence of any activity on its part (AW: 315–16).

Hobbes objected that the assumptions on which this whole way of thinking turned were fundamentally mistaken. The form (or essence, or nature) of a 'square' could not possibly produce effects in this way, or any way, because it was not a real entity—and only real entities produced effects. The words 'form', 'essence', and 'nature' and the rest signified nothing more than the way we think about a set of particular things that, in our experience, share certain properties. In *De corpore*, Hobbes concluded that this meant that formal causes, like final causes, were really efficient causes misdescribed, since they caused knowledge of the things that they signified, as knowing that something is a square is the cause of knowing that its internal angles are right angles (and vice versa). For Hobbes, to say that something was a square was already to acknowledge that it possessed the feature of right-angledness that was taken to be characteristic of things like *it*; but to say that its 'squareness' was a *cause* of its 'right-angledness' was at the best uninformative, for it was just to say that something was a square because it was a square (DCO: 132).

Hobbes's view was that essences were not queer metaphysical entities that were ontologically distinct and distinguishable from things themselves



but which, when combined with the relevant matter, brought something new into existence (DCO: 17). Rather, they were conceived only from sense perceptions of already-existing things and they persisted for only so long as the things themselves persisted: once the thing was gone, only the memory and the name remained, to remind us of what had been lost. Thus we arrive at the answer to the question of why generation cannot involve the communication of an essence. Essences are nothing other than definitions—words that specify the way we think about things which we take to embody certain properties. They do not precede existence and so cannot be transmitted to something that does not yet exist. Accordingly, they did not figure in Hobbes's definition of generation as a process by which something undergoes 'a change which makes us assign to, or remove from [it] the name that answers the question: "What is the thing?"' (AW: 58–59): for unlike Filmer's, his was a non-Aristotelian view of generation.

#### IV

So much is evident in their respective accounts of the generation of children. Where Filmer had attributed to the father a formal and efficient causal role in bringing the child into existence through generation, Hobbes spoke of the mother giving life to the child, 'not by *generation* but by *looking after* him', that is, by sustaining the life that distinguished it from a mass of matter (DC: 109; compare EL: 130–31).<sup>40</sup> Notice that on this account the giving of life and the preservation of life are all one: and that the child owes its life to whoever fulfils this function, for it is preserving life that gives 'the title to dominion' over the child. Again, on Hobbes's account the child was obliged to obey its parents not, as Filmer had claimed, because they 'begat him', but by its own consent, which for Hobbes was sufficiently expressed by the child accepting the nourishment offered by its mother (L: 139–40). Indeed, it owed its obedience to *anyone* who protected it and preserved its life—a point whose political resonance for Hobbes probably does not need to be emphasised.

What may require emphasis is that Hobbes understood the relation between subjects and the commonwealth on the same model. People gave life to the commonwealth not by generating it in Filmer's sense of that term—that is to say by communicating an essence via a particular transaction—but instead by looking after it: by preserving it in existence as it preserved them.<sup>41</sup> This meant preserving whichever individual or group of individuals bore the person of the sovereign, since strictly speaking an essence could not exist unless it was exemplified in an existent. As Hobbes put it in *Leviathan*, 'the Sovereign, is the publique Soule, giving Life and Motion to the Common-wealth; which expiring, the Members are governed by it no more, than the Carcasse of a man, by his departed Soule . . . [and] he that wants protection, may seek it anywhere; and when he hath it, is



obliged to protect his Protection as long as he is able' (L: 230).<sup>42</sup> That people did not transmit through generation the essence of the commonwealth, which was given life subsequently by their actions, meant that this essence did not lie in a specific type of polity as (in their different ways) Filmer and Smith had supposed. This is one reason why, for Hobbes, monarchy, aristocracy, and democracy were all valid ways of arranging a commonwealth (DC: 91–101).<sup>43</sup> At the same time, it suggested that, whatever those arrangements, it was for individuals to sustain the essence of the commonwealth in existence by acknowledging the sovereignty of a particular man or body of men and protecting and nourishing it as it protected them. To this end, they had to think of themselves not simply as servants but as active citizens, and act accordingly in order to keep the commonwealth going.<sup>44</sup>

There is another reason why Hobbes could envisage a gamut of valid institutional possibilities within the commonwealth to Filmer's one: his view of essences implied a scope and a content which Filmer's view could not admit. Hobbes's view stated that essences were words that specified how people conceived of things that, in their experience, shared certain properties. This in no way meant that people could *only* have conceptions which corresponded to things given in the external world, as it were, by nature: combine ideas derived from sense perceptions of a man and a horse and you arrive at the idea of a centaur (L: 16). The first two ideas Hobbes distinguished from the third as examples of simple and compounded imagination, respectively, because the first two involved imagining objects formerly presented to the senses as simple wholes—Bruce Forsyth and Red Rum, for instance—and the third involved combining these images together, perhaps with others, to produce something else again. This distinction was reflected in the characters of the words used to mark and signify these images.

In *De corpore* Hobbes divided words into seven categories, each of which incorporated other auxiliary classes of word. Of these categories and classes only three need be mentioned in this place. First, there were common names, which signified groups of simple ideas united by nature, whether man, horse, or tree. Second, there were universal names, each of which stood for one idea involved in all such unities, such as body, or brown, or square. Last, and most importantly for our purposes, there were abstract compounded names (DCO: 15–24). These were words which united common names and universal names. As such, they did not have a referent that existed in nature. Hobbes's position was that *no* abstract compounded names suggested specific ideas to the understanding. What they produced was images of past experience connected with those words and tintured by the affections and aversions of the person who read them or heard them (L: 30–31). This category included the names virtue, vice, justice, gravity, wisdom, and fear, besides (*pari passu*) sovereignty, protection, obedience, and commonwealth.

The centrality of such terms to the discussion of politics is almost axiomatic. But the obvious inference from Hobbes's philosophy of language is



that to use these terms is less to discuss ideas in the abstract than to raise compound images which touch the affections of the listener or reader. To do this can scarcely be thought to be a purely speculative activity because its effect is not simply cognitive, but practical: *ceteris paribus* it is not only to develop ideas but to affect conduct. It might be added that since Hobbes assimilated affections to conceptions, assimilated conceptions to images, and termed ‘the Imagination that is rayseed in man . . . by words’ the ‘understanding’, it could not be otherwise (L: 19). The question then becomes, with what arrangements would these words, and therefore pleasurable or frightful images, be connected?

It hardly need be said that this view of the understanding and its operations places speakers and writers in an extremely powerful position. For it is in their hands to connect names like ‘commonwealth’ and ‘protection’ which suggest pro-attitudes with arrangements of their own choosing (whether monarchy, aristocracy, or democracy), because these names did not imply only one set of conceptual contents. Indeed, they implied none at all in themselves—which is why Filmer’s qualification misses its mark.<sup>45</sup> At the same time, it placed a high degree of responsibility in their hands, for they had the power to direct people to the proper ends—or to other ends again.

Direction would be in some measure didactic, since it would involve expounding definitions designed to call to mind certain images. But not just *any* definitions would do: the relevant definitions could not be arbitrarily stipulated and then patterned in a self-consistent and orderly way. Hobbes worked on the assumption that they would need, as it were, to ‘hook on’ to experience, including previous linguistic experience, if they were to provide a linguistic context fit to guide listeners and readers to politically and morally desirable goals through the evocation of the painful or pleasurable images necessary to propel them towards those goals. It was only ‘from something pleasant’ or unpleasant that the thought of enjoying or avoiding it arose, remember; and ‘from the thought arises the notion of the path to secure it; and from the notion of the path arises the progression towards the object desired’. Thus he would claim in *Leviathan* to have ‘derived the Rights of Sovereigne Power, and the duty of Subjects hitherto, from the Principles of Nature onely; such as Experience has found true, or Consent (concerning the use of words) has made so’ (L: 255).<sup>46</sup> This claim was leading, perhaps, but no more so than any other: what mattered was the destination to which it led.

## V

The destination towards which Hobbes’s philosophical commitments pointed was a new account of the ‘Causes, Generation and Definition of a Common-wealth’ (L: 117). In that account, material, efficient, formal,



and final causality each had a role to play. The first part of *Leviathan* identified the material cause of the commonwealth, 'Man'. Man *qua* patient was acted upon by other objects in the external world impacting upon the organs of sense. These objects raised in him the desires necessary to produce the effect—the commonwealth—that Hobbes was investigating.

Hobbes identified one complex of desires in particular—men's 'fore-sight of their own preservation, and of a more contented life'—as the 'finall Cause, End or Designe' of the commonwealth (L: 117). Since, with him, final causes were really efficient causes that pushed people 'towards the object desired', it followed that men were both the agents and the patients and therewith the integral cause of the commonwealth—its 'Makers' and its 'Matter' (L: 221). Would it be right to think, then, that for Hobbes the commonwealth existed by necessity? There is a sense in which it would, but only the sense that a commonwealth either exists or it does not. If it exists, then it exists necessarily, because its integral cause must also exist; and such causes cause necessarily. If, however, it does not exist, then some cause *sine qua non* of its existence must be lacking. It may be, for instance, that men lose sight of their own preservation or fail to grasp the terms on which it is secured: if this happens, then the necessary conditions for the continuing existence of the commonwealth do not obtain, and it dissolves (L: 221–30).

Hobbes was therefore entirely serious when he wrote that a commonwealth without sovereignty was no commonwealth at all. A commonwealth could not exist without sovereignty, because sovereignty was its formal cause: the unity in which 'consisteth the Essence of the Commonwealth'. It was here that Hobbes, like Filmer, found it necessary to supplement causality with representation. Where Filmer had deployed the notion of representation to explain the unity of the entire human race in Adam's person, Hobbes deployed the same notion differently to explain a unity of another kind. Representation supplied what otherwise was missing from his account of the commonwealth: an account of its formal cause.

As we have seen, Hobbes's view was that formal causes were essences. But if, as he also thought, essences were definitions and formal causes were efficient causes, then, on the principle of transitivity, definitions were efficient causes. That is to say, they were causes not only of knowledge, as knowing the definition of a square causes us to know that its internal angles are right angles, but causes of action. Definitions evoked in the minds of those that read or heard them pleasurable or frightful images; properly ordered in combination with other definitions, they produced chains of thought that gave rise to actions of one sort or another. The right definitions pushed people in the right direction; the wrong definitions propelled them into absurdities of a very real kind. On this view, confusion over—or ignorance of—what sovereignty *is*, is less a forgivable failure of theorising than the prelude to the mortality of the commonwealth and, with this, of its several members.



On the one hand, this explains why Hobbes was concerned not only to see a high degree of control exerted over the promulgation of doctrines in the commonwealth (not least by having *Leviathan* taught in the universities) but also, and more importantly, to place the power of defining words into the sovereign's hands, and only his hands (L: 124–5; RC: 491). As *Elements of Law* put it, it was necessary 'there should be a common measure of all things that might fall into controversy; as . . . of what is to be called a right, what good, what virtue, what much, what little, what *meum* and *tuum*, what a pound, what a quart &c . . . This common measure [is not] to be found or known *in rerum naturâ* [and so] the reason of some man, or men, must supply the place thereof; and that man, or men, is he or they, that have the sovereign power' (EL: 180–1). Political authority and linguistic authority could not be separated; indeed, they stood and fell together, such that abrogating one's responsibilities in the one undermined the other.

On the other hand, it suggests that the exercise of linguistic authority is always to some extent an exercise in persuasion, whether it is conducted in a scientific mode or in a rhetorical mode; for it is always an attempt to raise in the minds of others the images sufficient to produce the conduct desired. Sometimes this will be achieved best by explaining the connexions of complex words to those whose understandings are equipped to grasp those connexions—perhaps because they were familiar with the specified conceptual content of those words and their connexions already (see DCO: 388)—at other times by the use of imagery with which pleasure or pain and danger are associated, at others again by the use of both together. Much would depend on the experiences of those whom one wished to persuade, a point which Hobbes acknowledged in his own practice when varying his mode of address in his political writings according to his intended audience.<sup>47</sup> His aim, in either case, was to modify the mental images with which these words were associated and to align word and image in ways that changed how people understood things and amended their conduct accordingly.

This aim bears upon the character of the argument of *Leviathan* in a profound way. For it suggests that the causal account of the generation of the commonwealth which it embodies should be read as it runs, that is, as an account of how the *name* 'commonwealth' is built up in logical sequence from other, successively more complex names, each of which has been settled unequivocally in its turn. Its purpose is to bring men to see the civil association of which they are a part in those terms and thence to play their parts in keeping it in existence. It is assuredly *not* what it has commonly been taken to be: namely a putative account of how the association to which Hobbes wished to attach that name is constructed in temporal sequence by individual agents.<sup>48</sup> For Hobbes's ambition was not 'the construction of political society from an ethical vacuum' but rather the reverse: to prevent its deconstruction *into* an ethical vacuum.<sup>49</sup> More, his was an attempt not only to show his readers the ease and very great danger of deconstructing it through their own misunderstanding, pride, and folly,



but to allay that danger. In redefining the commonwealth, Hobbes understood himself to be contributing one of the causal elements necessary to its continuing survival.

## NOTES

1. Abbreviations used: AW [*Anti-White*]: *Thomas White's De Mundo Examined*, edited by H. Whitmore Jones (London: Bradford University Press, 1976). BH: *Behemoth, or the Long Parliament*, edited by F. Tönnies, with an introduction by S. Holmes (Chicago: University of Chicago Press, 1990). DC [*De cive*]: *On the Citizen*, edited by R. Tuck, translated by M. Silverthorne (Cambridge: Cambridge University Press, 1998). DCO [*De corpore*]: *The English Works of Thomas Hobbes of Malmesbury*, volume 1, edited by W. Molesworth (London: John Bohn, 1839). EL: *The Elements of Law, Natural and Politic*, edited by J. C. A. Gaskin (Oxford: Oxford University Press, 1994). EW: *The English Works of Thomas Hobbes of Malmesbury*, in 11 volumes, edited by W. Molesworth (London: John Bohn, 1839–45). L: *Leviathan: or, the Matter, Forme, & Power of a Common-wealth Ecclesiasticall and Civil*, edited by R. Tuck (Cambridge: Cambridge University Press, 1996). RC: A Review, and Conclusion, to *Leviathan*. SL: *Six Lessons to the Professors of the Mathematics: The English Works of Thomas Hobbes of Malmesbury*, volume 7, edited by W. Molesworth (London: John Bohn, 1845). See in this connection Douglas Jesseph, "Hobbesian Mechanics", in *Oxford Studies in Early Modern Philosophy, Volume III*, edited by Daniel Garber and Steven Nadler (Oxford: Clarendon Press, 2006), 119–52.
2. For Hobbes as an enemy of Aristotelianism, see, e.g., Tom Sorell, *Hobbes* (London: Routledge & Kegan Paul, 1986), 2–3, 5, and Thomas A. Spragens, Jr., *The Politics of Motion: The World of Thomas Hobbes* (London: Croom Helm, 1973), 187–93.
3. Aristotle, *Metaphysics*, Book A, 980<sup>a</sup>–993<sup>a</sup> 25, in *The Basic Works of Aristotle*, edited by R. McKeon (New York: Random House, 1941), 689–712. Those criticized included Thales, Anaximenes, Diogenes, Anaxagoras, Empedocles, Leucippus, Democritus, Alcmaeon of Croton, Pythagoras and the Pythagoreans, Parmenides, Xenophanes, Cratylus, Heraclitus, and Plato.
4. For first philosophy in Hobbes, compare Jesseph, "Hobbesian Mechanics", *passim*, and Yves Charles Zarka, "First Philosophy and the Foundations of Knowledge", in *The Cambridge Companion to Hobbes*, edited by T. Sorell (Cambridge: Cambridge University Press, 1996), 62–85.
5. The honourable exception being Euclid in geometry.
6. See Hobbes, "A Minute or First Draft of the Optiques", excerpted in EW 7, 471: 'I [have] beene y<sup>e</sup> first to lay the grounds of two sciences; this of *Optiques* . . . and y<sup>t</sup> other of *Natural Justice*, which I have done in my booke DE CIVE'. Compare DCO, The Epistle Dedicatory, ix: 'Civil Philosophy [is] no older . . . than my own book *De Cive*'.
7. Thomas Smith, *De Republica Anglorum* (London, 1583), 3.
8. For fuller discussion, to which the present account is indebted, see Richard Tuck, "Hobbes and Democracy", in *Rethinking the Foundations of Modern Political Thought*, edited by A. Brett and J. Tully, with H. Hamilton-Bleakley (Cambridge: Cambridge University Press, 2006), 171–90.
9. I have amended Silverthorne's translation; for fear, compare L: 99.
10. For the classical notion, see Kurt von Fritz, *The Theory of the Mixed Constitution in Antiquity* (New York: Columbia University Press, 1954). For



- revivification, see, e.g., Philip Hunton, *A Treatise of Monarchie* (London, 1643), Henry Ferne, *Conscience Satisfied* (Oxford, 1643), 6. Compare Tuck, “Hobbes and Democracy”, 175–76.
11. Hobbes wrote elsewhere of ‘mixarchy’, which had been praised ‘by the name of mixed monarchy, though it were indeed nothing else but pure anarchy’. See B: 116–17.
  12. For opponents of the king, see, e.g., John Lilburne, *The Free-Man’s Freedome Vindicated* (London, 1647); Richard Overton, *An Arrow against All Tyrants* (London, 1646); William Prynne, *The Sovereigne Power of Parliaments and Kingdomes* (London, 1643); and, for discussion, Jonathan Scott, *Commonwealth Principles: Republican Writing of the English Revolution* (Cambridge: Cambridge University Press, 2004), 151–69, and Quentin Skinner, *Hobbes and Republican Liberty* (Cambridge: Cambridge University Press, 2008), 66–70, 84–85, 152–57.
  13. See variously Quentin Skinner, “The Ideological Context of Hobbes’s Political Thought”, *Historical Journal* 9, (1966): 286–317; *Reason and Rhetoric in the Philosophy of Hobbes* (Cambridge, 1996), 284–93; and *Republican Liberty*, 158–62; and compare Charles D. Tarlton, “The Despotical Doctrine of Hobbes, part ii”, *History of Political Thought*, 23 (2002): 62–89.
  14. Robert Filmer, *Observations Concerning the Originall of Government in Patriarcha and Other Writings*, edited by J. P. Sommerville (Cambridge: Cambridge University Press, 1991), 184. All subsequent works cited by Filmer are cited from this edition.
  15. Filmer, *Observations*, 186; Jean Bodin, *The Six Bookes of a Commonweale*, translated by. R. Knolles (London, 1606).
  16. For a selection of considerations of this sort, see Scott, *Commonwealth Principles*, passim, and compare Skinner, *Republican Liberty*, 198–210.
  17. As Ian Harris has demonstrated. Compare “Filmer”, in *The Continuum Companion to Locke* (London: Continuum, 2010), 57–61; “The Legacy of *Two Treatises of Government*”, *Eighteenth Century Thought*, 3 (2007): 143–67; and “Sex, Power and Consent in *Two Treatises of Government*”, unpublished paper, delivered 3 April 2004 to the British Society for the History of Philosophy Locke Tricentenary Conference, St. Anne’s College, Oxford. I am grateful to Dr. Harris for making available to me the typescript of his paper.
  18. Filmer, *Observations*, 187.
  19. Aristotle, *Generation of Animals*, 729<sup>a</sup> 28–730<sup>b</sup>, in *Works*, 676–78.
  20. Filmer, *Observations*, 192.
  21. See, e.g., David Hume, *A Treatise of Human Nature*, edited by D. Fate Norton and M. J. Norton (Oxford: Clarendon Press, 2000) 1.3.15, 116.
  22. Aristotle, *Physics* 194<sup>b</sup> 16–195<sup>a</sup> 3; and cf. *Metaphysics*, 983<sup>a</sup> 25–32, *Generation of Animals*, 715<sup>a</sup> 4–7, in *Works*, 240–41, 693, 665.
  23. Aristotle, *Generation of Animals*, 729<sup>b</sup> 7–8, 730<sup>b</sup> 3–4, in *Works*, 676, 678.
  24. Harris, “Filmer”, 57.
  25. See L: 466, ‘the Essence of a Man, which (they say) is his Soule’.
  26. Robert Filmer, *Patriarcha*, 6. See also Harris, “Sex, Power and Consent”.
  27. Harris, “Filmer”, 57, and compare idem., “Sex, Power and Consent”.
  28. Alexander Ross, *The First Booke of Questions and Answers upon Genesis* (London, 1620), 21.
  29. Harris, “Filmer”, 58. See *Patriarcha*, 11, 13, for Adam enjoying the ‘right and natural authority of a supreme father’ by ‘the law of nature’.
  30. Harris, “Filmer”, 58.
  31. Filmer, *Observations*, 187.
  32. Filmer, *The Anarchy of a Limited or Mixed Monarchy*, 138–9.



33. Edward Stillingfleet, "To Dr. Burthogge about Original Sin, and the Covenant with Adam, and the meaning of Gen 3. 15", in *Miscellaneous Discourses on Several Occasions*, edited by J. Stillingfleet (London, 1735), 346–56, at 356. The logic of representation could also be used, as it was by Stillingfleet himself, to explain why all of mankind had sinned in the sin of Adam. For discussion, see Ian Harris, "The Politics of Christianity", in *Locke's Philosophy: Content and Context*, edited by G. A. J. Rogers (Oxford: Clarendon Press, 1994), 197–215, esp. 198–212.
34. Thus Filmer could write of the 'fatherly right of sovereign authority': *Patriarcha*, 11.
35. Filmer, *Patriarcha*, 24.
36. Filmer, *Patriarcha*, 14, alluding to Plato, *Statesman*, 259<sup>b</sup>.
37. See, e.g., David Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes* (Oxford: Clarendon Press, 1969); Jean Hampton, *Hobbes and the Social Contract Tradition* (Cambridge: Cambridge University Press, 1986); Gregory S. Kavka, *Hobbesian Moral and Political Theory* (Princeton, NJ: Princeton University Press, 1986); Jody S. Kraus, *The Limits of Hobbesian Contractarianism* (Cambridge: Cambridge University Press, 1993).
38. As the subtitle of *Leviathan—The Matter, Form & Power of a Commonwealth Ecclesiasticall and Civil*—indicates.
39. This is a point about scope, not content. Hobbes did not construe these properties in remotely the same terms as Aristotle (and Filmer) did. For an inventory of differences, see Kenneth C. Clatterbaugh, *The Causation Debate in Modern Philosophy, 1637–1739* (New York, Routledge, 1999), 1–16, 69–81.
40. Here if anywhere is the place to observe that on Hobbes's account the man and woman together formed an integral cause of their offspring; but (*pace* Filmer) it might just as well be the woman as the man that was the agent in generation, depending upon which raised the desire to procreate in the other. This would be a matter for tentative empirical investigation.
41. For an alternative interpretation, which presents Hobbes as following out in his account of the generation of the state a metaphor of procreation on the model presumed by Filmer, see Quentin Skinner, "Hobbes on Person, Authors and Representatives", in *The Cambridge Companion to Hobbes's Leviathan*, edited by P. Springborg (Cambridge: Cambridge University Press, 2007), 157–80.
42. A point which Hobbes underlined by adding in RC (484) a further law of nature to the nineteen previously identified, 'That every man is bound . . . as much as in him lieth, to protect in Warre, the Authority, by which he is himself protected in time of Peace'.
43. Which is not to say that he thought them all equally advantageous: see DC: 115–26, esp. 117, 'Monarchy Is the Best of the Listed Kinds of Commonwealth'.
44. See L: 174–76 for the various ways in which their public activity nourishes the commonwealth.
45. The present argument tends to confirm the claim recently made by Quentin Skinner that Filmer failed to understand that Hobbes was trying to persuade people that 'absolute monarchies may be no less deserving of the name of commonwealths than the freest and most democratic of free states', but provides an alternative explanation of that failure and its grounds. See Skinner, *Republican Liberty*, 209–10.
46. See also EL: 41. Concordant claims appeared in Hobbes's other scientific writings. See, e.g., DCO: 531.



47. Compare the scientific mode of presentation in the *Elements of Law*, circulated in manuscript form to a select audience, and the combination of scientific and rhetorical modes in *Leviathan*, which was intended for a far wider audience. For which, see G. Vaughan, “The Audience of *Leviathan* and the Audience of Hobbes’s Political Philosophy”, *History of Political Thought*, 22, 3 (2001): 448–71.
48. Neither were there good grounds for supposing that the order of development of the one would track the other. See DCO: 24: ‘in the mind one idea or phantasm succeeds another, and to this a third; so to one name is added another and another successively, and of all of them is made one *compounded* name. Nevertheless we must not think bodies which are without the mind, are compounded in the same manner . . . though such has been the philosophy of many’. It remains so.
49. John Dunn, *The Political Thought of John Locke* (Cambridge: Cambridge University Press, 1969), 79.



# 7 Hume, Causal Realism, and Free Will

*Peter Millican*

My aim in this chapter is to present what I consider to be the decisive objection against the ‘New Hume’ causal realist interpretation of Hume, and to refute three recent attempts to answer this objection. I start in §I with an outline of the ‘Old’ and ‘New’ interpretations. Then §II sketches the traditional case in favour of the former, while §III presents the decisive objection to the latter, based on Hume’s discussions of ‘Liberty and Necessity’ (i.e., free will and determinism). In §§IV–VI, I consider in turn the recent responses of Helen Beebe, Peter Kail, and John Wright, and explain why these fail. My conclusion in §VII is that the New Hume can reasonably be considered as refuted, unless and until a more successful response is forthcoming, which (to me at least) looks extremely unlikely.

## I. THE OLD AND THE NEW HUME

David Hume is universally associated with the *regularity theory of causation*, which is generally understood as involving a ‘reduction’ of causal relations between objects to regular succession (and a corresponding association of ideas in the observing mind) through his two definitions of cause:

THERE may two definitions be given of this relation . . . We may define a CAUSE to be ‘An object precedent and contiguous to another, and where all the objects resembling the former are plac’d in like relations of precedency and contiguity to those objects, that resemble the latter.’ If this definition be esteem’d defective, because drawn from objects foreign to the cause, we may substitute this other definition in its place, *viz.* ‘A CAUSE is an object precedent and contiguous to another, and so united with it, that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other.’ (*T* 1.3.14.31, cf. *E* 7.29)<sup>1</sup>

Interpreted as the traditional ‘Old Hume’, he is denying that there is anything more to causation in objects than is expressed in the two definitions,



thus rejecting, in recent parlance, any ‘thick connexions’ between objects.<sup>2</sup> Or to adopt a useful convention introduced by Galen Strawson, the regularity theory denies that there is any such thing as *Causation* (with an upper-case ‘C’):

For present purposes . . . ‘Causation’ may be merely negatively defined: it covers any essentially *non-Regularity-theory* conception of causation. More positively . . . to believe that causation is in fact Causation is simply to believe (A) that there is something about the fundamental nature of the world in virtue of which the world is regular in its behaviour; and (B) that that something is what causation is, or rather is at least an essential part of what causation is or involves.<sup>3</sup>

Strawson himself insists that Hume is—contrary to the dominant tradition—a true believer in such metaphysically heavyweight *Causation*.<sup>4</sup> An early version of such a reading can be found in the work of Norman Kemp Smith,<sup>5</sup> but its first systematic presentation was in John Wright’s 1983 book which gave it a name: *The Sceptical Realism of David Hume*. ‘Sceptical realism’ as understood by Wright combines epistemological and conceptual modesty with natural judgement:

In spite of the fact that real causal forces in nature are inconceivable to us, we judge that these forces exist. . . . Reason leads us to the conclusion that our ideas of cause and effect are distinct. Yet natural instinct leads us to a directly contrary conclusion: namely, that there is an objective necessary connection relating those objects which we experience as constantly conjoined. . . . The resolution of the conflict between the conclusions of instinct and reason is provided by a mitigated scepticism which recognises the inadequacy of our ideas of objects and yet ascribes causal necessity to the external objects themselves on the basis of the criterion provided by the natural instinct.<sup>6</sup>

One of the strengths of Wright’s interpretation is the similarity that he posits between Hume’s views on causation and on the external world:<sup>7</sup> thus Humean ‘sceptical realism’—in Wright’s sense—deserves to be understood in this broad manner. The narrower claim that Hume is specifically a (capital ‘C’) *Causal* realist acquired a distinctive nickname in 1991 when strongly criticised in Ken Winkler’s paper “The New Hume”, whose title implicitly recognised the significance of the novel interpretative trend he was attacking. By this time, Wright had been joined not only by Strawson but also by Donald Livingston, Janet Broughton, Edward Craig, and Michael Costa, to be followed in due course by John Yolton, Stephen Buckle, and Peter Kail.<sup>8</sup> This impressive lineup, all ranged in radical opposition to the ‘Old’ Humean stereotype, seemed fully to vindicate the aptness of Winkler’s ‘New Hume’ nickname.



But it would be wrong to assume that all these ‘New Humean’ accounts are essentially the same. As Kail observes, ‘What unites these readings is simply the *rejection* of a positive regularity reading of Hume and nothing more’.<sup>9</sup> Kail himself, unlike Wright and Strawson, is reluctant to ascribe to Hume a definite belief in Causation, suggesting a more agnostic position. But he nevertheless clearly aligns himself with what he calls this ‘revisionist’ camp, even while questioning whether Hume is a committed Causal realist. This might seem puzzling, but his point is that the central question is not whether Hume has a firm commitment to Causal powers, but rather, whether Hume’s epistemology—in particular his theory of ideas—leaves any *possible* room for them:

The revisionists’ target was a Hume who held a regularity theory of the metaphysics of causation. [Wright, Craig and Strawson] were further united in the conviction that the metaphysical doctrine was supposed to follow from semantic premises. Hume attempts to show that the notion of ‘necessary connection in the objects’ lacks any meaning, and so the issue of whether there is any ‘in the objects’ cannot even be intelligibly raised. . . . The real debate is whether we should maintain the standard reading, and reject, or reinterpret, the apparent references to hidden powers because of the alleged strictures of the theory of ideas or think that the presence of hidden power talk suggests that the cognitive strictures of the theory of ideas are not quite what they seem.<sup>10</sup>

On the fundamental issue, this seems right:<sup>11</sup> the characteristic ‘Old Hume’ position rules out any notion of (thick, upper-case) Causal powers as unintelligible, on the ground that there is no impression source for any such idea. Hence interpreting Hume as agnostic about such powers—at least if this is taken to imply that he considers their existence a meaningful possibility—should count as a ‘New Humean’ rather than an ‘Old Humean’ (or neutral) position.

In this chapter,<sup>12</sup> I shall follow Kail’s understanding of ‘the real debate’ as hinging on the question of whether or not Hume considers the notion of thick Causal powers to be *meaningful* or *intelligible*, rather than whether or not he is a committed *realist* about such powers. And accordingly, I shall generally refer to the revisionist interpretation using Winkler’s convenient ‘New Hume’ nickname rather than the potentially misleading terms ‘Causal *realism*’ or ‘sceptical *realism*’. Sometimes in the past, a failure to draw the distinction clearly in this way has led to confusion, as for example in the introduction to Winkler’s own eponymous paper:

I will argue that Hume *refrains from affirming* that there is something in virtue of which the world is regular in the way it is. This is not to *deny* that there is such a thing, but merely not to believe in it. Defenders of the New Hume sometimes ease their task by supposing that according to the



standard view, Hume positively denies the existence of secret powers or connections. They argue (rightly, in my view) that a positive denial runs counter to Hume's scepticism. But a refusal to affirm such powers or connections suits Hume's scepticism perfectly . . . <sup>13</sup>

If the 'powers or connections' being alluded to here are understood as *thick* powers or connections, going beyond Hume's two definitions, then, as Kail points out, ' "that . . . Hume positively denies the existence of [such] powers or connections" . . . *was* and *is* the standard view'.<sup>14</sup> So here Winkler seems to be suggesting that Hume is agnostic about Causation in the way that Kail claims for New Humeanism. Fortunately §2 of Winkler's paper—on 'The scope (or force) of the theory of ideas'—remedies the confusion:

[We are] free to suppose that Hume's scepticism consists in a refusal to affirm the existence of Causation, a refusal rooted in the belief that there is no notion of Causation to be affirmed (or denied, or even entertained as a possibility). The alleged notion of Causation is (to borrow from *Enquiry* §12) a notion so imperfect 'that no sceptic will think it worth while to contend against it' [E 12.16].<sup>15</sup>

Now it becomes clear that Winkler is indeed an 'Old Humean' of the broadly traditional sort, though his view remains more nuanced than this crude categorisation might suggest.<sup>16</sup>

## II. THE STANDARD OLD HUME READING

The traditional interpretation of Hume is based mainly on his well-known argument concerning 'The Idea of Necessary Connexion', as presented in *Treatise* 1.3.14 and *Enquiry* 7, and nicely summarised at the beginning of the *Treatise* version:

*What is our idea of necessity, when we say that two objects are necessarily connected together?* Upon this head I repeat what I have often had occasion to observe, that as we have no idea, that is not deriv'd from an impression, we must find some impression, that gives rise to this idea of necessity, if we assert we have really such an idea. In order to this I consider, in what objects necessity is commonly suppos'd to lie; and finding that it is always ascrib'd to causes and effects, I turn my eye to two objects suppos'd to be plac'd in that relation; and examine them in all the situations, of which they are susceptible. I immediately perceive, that they are *contiguous* in time and place, and that the object we call cause *precedes* the other we call effect. In no one instance can I go any farther, nor is it possible for me to discover any third relation betwixt these objects. I therefore enlarge my view to comprehend



several instances; where I find like objects always existing in like relations of contiguity and succession. At first sight this seems to serve but little to my purpose. The reflection on several instances only repeats the same objects; and therefore can never give rise to a new idea. But upon farther enquiry I find, that the repetition is not in every particular the same, but produces a new impression, and by that means the idea, which I at present examine. For after a frequent repetition, I find, that upon the appearance of one of the objects, the mind is *determin'd* by custom to consider its usual attendant, and to consider it in a stronger light upon account of its relation to the first object. 'Tis this impression, then, or *determination*, which affords me the idea of necessity. (T 1.3.14.1)

There are many passages in the *Treatise*, the *Enquiry*, and the summary in the *Abstract* that seem to confirm Hume's aim as being to establish the *meaning* of our attributions of necessity or causal power, through the identification of the source impression, for example:<sup>17</sup>

Necessity, then, . . . is nothing but an internal impression of the mind . . . Without considering it in this view, we can never arrive at the most distant notion of it, or be able to attribute it either to external or internal objects. . . . (T 1.3.14.20)

The question is, what idea is annex'd to these terms [*power*, or *force*, or *energy*]? . . . Upon the whole, . . . either we have no idea at all of force and energy, and these words are altogether insignificant, or they can mean nothing but that determination of the thought, acquir'd by habit, to pass from the cause to its usual effect. (A 26)

There are no ideas, which occur in metaphysics, more obscure and uncertain, than those of *power*, *force*, *energy*, or *necessary connexion* . . . We shall, therefore, endeavour, in this section, to fix, if possible, the precise meaning of these terms . . . all our ideas are nothing but copies of our impressions, or, in other words, . . . it is impossible for us to *think* of any thing, which we have not antecedently *felt*, either by our external or internal senses. . . . this customary transition of the imagination from one object to its usual attendant, is the sentiment or impression, from which we form the idea of power or necessary connexion. . . . When we say . . . that one object is connected with another, we mean only, that they have acquired a connexion in our thought, and give rise to this inference, by which they become proofs of each other's existence. . . . (E 7.3, 7.4, 7.28)

Moreover, the *Treatise* and *Enquiry* discussions both culminate with Hume's two 'definitions of cause' (T 1.3.14.31, as quoted earlier, and



*E* 7.29), which again seems to confirm that the ultimate aim of Hume's quest for the impression of necessary connexion is the clarification of meanings. If this is the case, then the result of that quest would seem to imply a constraint on what we can mean by 'necessary connexion', thus giving rise to the Old Hume interpretation.

### III. OF LIBERTY AND NECESSITY

All this is very familiar, though the interpretation of the sections on necessary connexion has been subject to considerable debate, which I do not propose to add to here.<sup>18</sup> Instead, I want to move forward immediately to Hume's treatment 'Of Liberty and Necessity' (*Treatise* 2.3.1–2 and *Enquiry* 8), which contains the main application of his two definitions.<sup>19</sup> Here he appeals to them in his main argument to establish 'the doctrine of necessity', that is, the doctrine that determinism applies to human actions and the mind's operations, just as it does to material things. This part of Hume's discussion is very similar in both works, but here I shall focus mainly on the *Enquiry*, since this is appealed to as the authoritative—or at least more clearly Causal realist—source by New Humeans.<sup>20</sup> Following each *Enquiry* quotation, however, I shall also cite the parallel *Treatise* passage for reference.<sup>21</sup>

Hume starts his argument for 'the doctrine of necessity' by focusing on our understanding of necessity as we attribute it to matter:

It is universally allowed, that matter, in all its operations, is actuated by a necessary force, and that every natural effect is so precisely determined by the energy of its cause, that no other effect, in such particular circumstances, could possibly have resulted from it. . . . Would we, therefore, form a just and precise idea of *necessity*, we must consider whence that idea arises, when we apply it to the operation of bodies. (*E* 8.4, cf. *T* 2.3.1.3)

He then refers back to his two definitions of *cause*, as set out at *E* 7.29, and uses these to characterise *necessity* in an exactly corresponding way, drawing the obvious moral for how its presence is to be identified in human actions:

These two circumstances form the whole of that necessity, which we ascribe to matter. Beyond the constant *conjunction* of similar objects, and the consequent *inference* from one to the other, we have no notion of any necessity, or connexion. If it appear, therefore, that all mankind have ever allowed . . . that these two circumstances take place in the voluntary actions of men, and in the operations of mind; it must follow, that all mankind have ever agreed in the doctrine of necessity. . . . (*E* 8.5–6, cf. *T* 2.3.1.4)

Having set this agenda, Hume devotes the next fourteen paragraphs (*E* 8.7–20) to arguing at length, and with a wide range of illustrative examples,



that human actions do indeed manifest uniformity to a similar extent to what we observe in the material world, and that this uniformity is generally recognised and taken for granted as a basis for inductive prediction. The following passage sums up these two claims, and draws the desired conclusion—that insofar as there is any substance to the issue, the doctrine of necessity is implicitly accepted by ‘all mankind’, even if most are reluctant to acknowledge this in so many words:

Thus it appears, not only that the conjunction between motives and voluntary actions is as regular and uniform, as that between the cause and effect in any part of nature; but also that this regular conjunction has been universally acknowledged among mankind . . . this experienced uniformity in human actions is a source, whence we draw *inferences* concerning them . . . [Such] inference and reasoning concerning the actions of others enters so much into human life, that no man, while awake, is ever a moment without employing it. Have we not reason, therefore, to affirm, that all mankind have always agreed in the doctrine of necessity, according to the foregoing definition and explication of it? (*E* 8.16–17, cf. *T* 2.3.1.16–17)

This essentially completes the main argument: Hume takes himself to have shown that the two definitional criteria for ascribing necessity are both fulfilled by human actions, and that these characteristics of actions are generally recognised.

Hume remarks, however, that this conclusion raises an obvious puzzle as to why so many people who ‘have ever . . . acknowledged the doctrine of necessity, in their whole practice and reasoning’, are so reluctant ‘to acknowledge it in words’ (*E* 8.21).<sup>22</sup> The answer, Hume suggests, lies in two complementary errors—people imagine that they detect ‘something like a necessary connexion’ in the operations of matter, and also suppose that they can feel the *absence* of any such connexion in the operations of mind:

Men still entertain a strong propensity to believe, that they penetrate farther into the powers of nature, and perceive something like a necessary connexion between the cause and the effect. When again they turn their reflections towards the operations of their own minds, and *feel* no such connexion of the motive and the action; they are thence apt to suppose, that there is a difference between the effects, which result from material force, and those which arise from thought and intelligence. (*E* 8.21)

Their ‘strong propensity to believe’ that they can ‘penetrate . . . into the powers of nature’ naturally leads philosophers to think that *genuine* necessity—of the sort that supposedly applies to bodies—must involve something more than mere constant conjunction and inference. But such thinking, Hume insists, can be quickly refuted:



It may . . . perhaps, be pretended, that the mind can perceive, in the operations of matter, some farther connexion between the cause and effect; and a connexion that has not place in the voluntary actions of intelligent beings. . . . [However] . . . a constant conjunction of objects, and subsequent inference of the mind from one to another . . . form, in reality, the whole of that necessity, which we conceive in matter . . . [if we] suppose, *that we have some farther idea of necessity and causation* in the operations of external objects . . . there is no possibility of bringing the question to any determinate issue, while we proceed upon so erroneous a supposition. (E 8.21-22, my emphasis)

This same crucial point—focusing on the absence of any ‘farther idea of necessity’—is emphasised pithily in the short final paragraph of the summary discussion in the *Abstract*:

Our author pretends, that this reasoning puts the whole controversy in a new light, by giving a new definition of necessity. And, indeed, the most zealous advocates for free-will must allow this union and inference with regard to human actions. They will only deny, that this makes the whole of necessity. But then they must shew, that we have an idea of something else in the actions of matter; which, according to the foregoing reasoning, is impossible. (A 34, my emphasis)

Having settled the issue of necessity, the *Enquiry* discussion quickly moves on to the second stage of Hume’s ‘reconciling project’ by considering ‘what is meant by liberty, when applied to voluntary actions’ (E 8.23).<sup>23</sup> Just as he has shed light on ‘the question of liberty and necessity’ with his ‘new definition of necessity’, so he now proceeds to give a new ‘definition . . . of liberty’ (E 8.24):

By liberty, then, we can only mean *a power of acting or not acting, according to the determinations of the will*; that is, if we chuse to remain at rest, we may; if we chuse to move, we also may. (E 8.23)

Hume ends *Enquiry* 8 Part 1 by stressing ‘the advantage of definitions’ (E 8.25), mentioning both his definitions of cause (informed by his account of ‘the origin of the idea . . . [of] *necessary connexion*’), and also his ‘definition above mentioned . . . [of] liberty’. Just as in the *Abstract*, it is quite clear that he sees his definitions of the relevant terms as providing his key novel contribution to the discussion.

Having finished his theoretical argument, in *Enquiry* 8 Part 2 Hume turns to address the practical accusation that his views have ‘dangerous consequences to religion and morality’ (E 8.26, T 2.3.2.3). The first paragraph of his answer is copied largely verbatim from the *Treatise*:



Necessity may be defined two ways, conformably to the two definitions of *cause*, of which it makes an essential part. It consists either in the constant conjunction of like objects, or in the inference of the understanding from one object to another. Now necessity, in both these senses . . . has universally, though tacitly, in the schools, in the pulpit, and in common life, been allowed to belong to the will of man; and no one has ever pretended to deny, that we can draw inferences concerning human actions, and that those inferences are founded on the experienced union of like actions, with like motives, inclinations, and circumstances. The only particular, in which any one can differ, is, that either, perhaps, he will refuse to give the name of necessity to this property of human actions: But as long as the meaning is understood, I hope the word can do no harm: Or that *he will maintain it possible to discover something farther in the operations of matter*. But this, it must be acknowledged, can be of no consequence to morality or religion, whatever it may be to natural philosophy or metaphysics. *We may here be mistaken in asserting, that there is no idea of any other necessity or connexion in the actions of body*: But surely we ascribe nothing to the actions of the mind, but what every one does, and must readily allow of. *We change no circumstance in the received orthodox system with regard to the will, but only in that with regard to material objects and causes*. Nothing therefore can be more innocent, at least, than this doctrine. (E 8.27, emphasis added, cf. T 2.3.2.4)

Hume's strategy here is very clear, and entirely in line with what has gone before. His response to the imagined objection is to run through his main argument, and to draw attention to the most likely source of disagreement, namely, that his opponent 'will maintain it possible to discover something farther in the operations of matter'. He then alludes to his earlier answer to this disagreement (cf. the earlier quotation from E 8.21–2): his assertion 'that there is no idea of any other necessity or connexion in the actions of body'. But while making clear that this is his answer—and without in any way withdrawing it or suggesting that it is inadequate—he goes on to provide an *additional* consideration that can be invoked *even if* that assertion 'may here be mistaken'.<sup>24</sup> Suppose that it is mistaken, and that we can indeed form an idea of some stronger type of necessity in matter. Nevertheless, *Hume points out that his mistake would then concern what he ascribes to matter, not what he ascribes to the mind*. So even if his assertion 'that there is no idea of any other necessity or connexion in the actions of body' is wrong, he cannot here be criticised on *moral* or *religious* grounds, because morality and religion are concerned with the nature of humanity, not the nature of matter, and he 'change[s] no circumstance in the received orthodox system with regard to the [human] will'.

Note, however, the very clear implication of this paragraph—following exactly in the spirit of the preceding argument—that Hume disagrees



with ‘the received orthodox system . . . with regard to material objects and causes’, and does so precisely by rejecting the ‘erroneous supposition . . . that we have some farther idea of necessity and causation in the operations of external objects’ (E 8.22). Hume’s distinctive position, in other words, is that we cannot even conceive of any type of ‘necessity’ or ‘causation’ that goes beyond the bounds of his two definitions. His imagined opponent purports to have such a conception, and to attribute it to bodies, ‘denying that [the definitions] make the whole of necessity’ (A 34) and ‘maintain[ing that] there is something else in the operations of matter’ (T 2.3.2.4). If this opponent were correct, Hume clearly implies, he himself would be ‘mistaken’, so his own position must be that his two definitions *do* ‘make the whole of necessity’ and that *there is nothing else* [to necessity] ‘in the operations of matter’. His ground for asserting this is very straightforward and entirely consistent in the *Treatise*, the *Abstract*, and the *Enquiry*: it is simply to insist against his opponent *that we have no such idea*, and hence that the attribution cannot be made.

The relevance of all this to the New Hume debate is equally straightforward and obvious. For *the New Humean position is precisely that of Hume’s opponent who claims that there is something more to ‘genuine necessity’ than is captured by Hume’s two definitions*. Hume takes himself to have a quick and decisive answer to this claim, in denying that there can be any such conception. Thus Hume’s main argument concerning ‘liberty and necessity’ runs directly contrary to the New Humeans’ position. He is here denying exactly what they assert, namely, that we can coherently ascribe to things some kind of ‘upper-case’ Causation or ‘thick’ necessity that goes beyond his two definitions. If we could indeed do this, then his imagined opponent would be able to ascribe that thick necessity to matter but not to minds, and thus undermine Hume’s claim of equivalence between the necessity of the two domains, which is the entire point of his argument. Nor can there be any serious doubt about his intentions here, for the same argument occurs in the *Treatise*, the *Abstract*, and the *Enquiry*, and it is the principal application of his two definitions in all three of these works. Those definitions are clearly intended precisely for this role, and it is a role that requires them to be interpreted *semantically* rather than merely *epistemologically*: as constraining what we are able to *think* or *mean* or coherently refer to. Here, then, we seem to have exactly the kind of argument which in §I we took to characterise the Old Humean position: an argument denying ‘thick’ Causal powers in objects, on the basis that any term that purports to refer to such powers ‘lacks any meaning, and so the issue of whether there is any “in the objects” cannot even be intelligibly raised’.<sup>25</sup>

Hume’s application of his definitions of cause to the ‘doctrine of necessity’ is not particularly subtle or complex, and it is very explicit. Moreover, in the *Enquiry* the definitions occur at E 7.29, and their application starts at E 8.5, just six paragraphs apart in adjacent sections whose titles are



clearly related ('Of the Idea of Necessary Connexion' and 'Of Liberty and Necessity'). So it is surprising how widely this link has been ignored in discussions of his philosophy, not least within the New Hume debate.<sup>26</sup> Very recently, however, there have been three discussions that (either explicitly or implicitly) contest my claim that 'Of Liberty and Necessity' provides crucial evidence against the New Hume, from Helen Beebee, Peter Kail, and John Wright.<sup>27</sup> Let us examine these in turn.

#### IV. BEEBEE ON LIBERTY AND NECESSITY

Helen Beebee sets out to argue that 'Hume's discussion of free will provides virtually no additional evidence, let alone decisive evidence, either for the traditional interpretation to which Millican subscribes or for any other'.<sup>28</sup> In particular, therefore, she attempts to reconcile that discussion with the New Humean 'sceptical realist' reading, which she characterises as follows:

The *sceptical realist interpretation* . . . casts Hume as a firm believer in real causal powers, and takes Hume to think that these powers are what our ordinary causal thought and talk refer to. A central feature of the sceptical realist interpretation is the claim that Hume's primary point in his discussion of causation is an epistemological one. While our habits of expectation generate belief in real powers—when the transition in the mind from cause to effect generates belief *that* the first event causes the second, that belief really is a belief about the existence of a real power—we can never come to grasp the *nature* of that power, since our idea of it is generated not by the power itself but by the felt transition of the mind. So it makes sense to believe in real powers—indeed, belief in them is mandatory because it arises as a result of natural processes in the imagination—despite the fact that our *idea* of those powers is deficient: we cannot, as Strawson (1989: 127) puts it, form a 'positively or descriptively contentful conception' of them. (pp. 415–6)

There are points at which this characterisation of 'sceptical realism' could be challenged, and (as we saw in §I) not all New Humeans would agree with the general claim that Hume is 'a firm believer' in thick connexions.<sup>29</sup> But in fact this is a central aspect of Beebee's favoured brand of sceptical realism, which explains the Causal realist commitment as a 'natural belief: . . . one that is forced upon us by the operations of the imagination' (p. 428).<sup>30</sup> On her approach, the impression of necessary connexion, though itself arising internally from 'the felt transition of the mind' when we make a causal inference, naturally and irresistibly *represents* something quite different, namely, the supposed 'real power' in the objects themselves.<sup>31</sup> Thus the subjective character of the impression—and hence the corresponding idea—is sharply distinguished from their objective content. Indeed, this



sharp distinction is key to Beebee's reconciliation of sceptical realism (so understood) with Hume's argument concerning liberty and necessity. For as we have seen, the crucial move in that argument is Hume's repeated insistence that our understanding of necessity is constrained by our idea of it, and that we have no 'farther idea of necessity and causation in the operations of external objects' (E 8.22). But if our one *legitimate* idea of necessity *already* makes reference to the sceptical realist's supposed thick causal powers (despite that idea's subjective origin), then this constraint on our understanding need not apparently be any obstacle to a uniformly New Humean understanding of necessity.

We shall come back to the question of whether this distinction between an idea's subjective character and its representative content is plausibly Humean, but for the moment let us allow it. Beebee's sceptical realist still faces the challenge of showing that such a distinction is at work in Hume's discussion of necessary connexion, which seems highly focused on identifying and clarifying the circumstances and character of the relevant *subjective* impression, and *defining* accordingly what it is to be a cause. This procedure sits uneasily with the suggestion that Hume takes the impression to be making reference beyond, to a supposed objective thick power that outruns the definitions. Moreover, any such further reference seems to play no role when he comes to apply his definitions to the question of liberty and necessity, an argument which can be crudely represented as follows:

- Def* Necessity is to be defined in terms of constant conjunction and inference only.
- CCI* Constant conjunction and inference apply just as much to the moral as the physical world, and are universally recognised as doing so.
- Nec* Therefore necessity applies as much to the moral as the physical world. It is not possible to maintain that there is a thick necessity 'in the operations of matter' which is not present 'in the voluntary actions of intelligent beings'.

On the Old Humean view, the kind of 'definition' involved in *Def* is semantic—specifying the *meaning* of 'necessity'—thus making the argument very straightforward. Having acknowledged this (p. 424), Beebee considers whether other readings might enable the inference to go through:

One way in which one might try to proceed would be to claim that Hume is making an epistemic point: since our *grounds* for believing in thick necessity in both the human and non-human cases are the same, the libertarian has no right to claim that thick necessity is present in the first case and absent in the second. But this, just by itself, is not good enough: Hume's argument is not that (as far as our best evidence tells us) necessity is *in fact* present in both cases; it is that *everyone agrees* (on reflection, and once they have accepted the two definitions) that



this is so. . . . The sceptical realist interpreter thus needs to square the claim that the two definitions do not exhaust the nature of necessity with Hume's claim that, once we accept the two definitions, we will all *in fact* agree on the doctrine of necessity—and not merely with the claim that there will be no empirical *grounds* for disagreement. (p. 425)

Beebee is right to rule out the 'mere empirical grounds' reading: Hume is not simply arguing that—given acceptance of his definitions—*CCI* confirms *Nec* empirically. But her gloss on the argument's force does not go far enough in limiting it to *universal reflective agreement* on *Nec* subject to *CCI*. Hume's words suggest something even stronger: that once the two definitions are agreed, *there is no conceptual space left* for acknowledging *CCI* while denying *Nec*.<sup>32</sup>

Leaving this reservation aside, Beebee now goes on to observe that 'Hume's epistemology, as far as causation is concerned, has both naturalistic and normative aspects'. Part of the naturalistic side is that we find ourselves making causal judgements when the impression of necessary connexion arises in inductive inferences, and this occurs when we make predictions about the behaviour of people as well as things. As for the normative side:

we should not merely restrict our belief in causation to those cases where the imagination happens to deliver the impression of necessary connection: Hume clearly tells us, in his 'rules by which to judge of causes and effects' . . . that we *ought* to seek out hidden causes, for example. One of the aims of Hume's discussion of the doctrine of necessity is to show that we do, in fact, subscribe to Hume's rules in the human case every bit as much as in the non-human case; . . . (p. 426)

It is this normative aspect of Hume's theory to which Beebee appeals as a way of generalising 'the belief in thick necessity [which] is delivered by constant conjunction and the felt determination of the mind' (p. 427). Hume's rules commit us normatively to extending such belief from the causal interactions that have directly prompted our customary inferences to other interactions that are appropriately similar. To summarise:

Sceptical realist interpreters claim that belief in real powers is a natural belief: it is one that is forced upon us by the operations of the imagination. And [given] . . . the normative aspect of Hume's epistemology of causation . . . , Hume holds that that belief is one that we are *in fact*, whether we like it or not, committed to in cases where, for example, we believe that apparent irregularities in human behaviour are explained by hidden differences in character or motives. So . . . Hume's argument shows that . . . the necessity [the libertarian] is in fact committed to in the human realm just *is* the same as the necessity she is committed to in the non-human realm. (p. 428)



Beebee concludes by giving a very general gloss on Hume's two definitions which ties in neatly with this account (pp. 429–30).<sup>33</sup> The second definition focuses on the impression of necessary connexion, which fixes the meaning of our causal terms (either because our idea is a straightforward copy of the impression, as on the Old Hume account, or because it represents a thick power, on Beebee's favoured New Humean reading). Meanwhile, the first definition specifies the circumstances under which 'causal talk is appropriate' more widely, even in cases where no such impression arises.

Overall, Beebee has devised an ingenious story for reconciling Hume's argument concerning liberty and necessity with the principles of her Causal realist New Hume. But unfortunately, it cannot stand up to close critical examination. To start with her last point, there is no textual evidence that Hume takes his argument to be turning on the sorts of normative considerations that her account would imply. On the contrary, as we saw earlier in §III, Hume consistently points to issues of *meaning* and *the limits of our ideas* as the locus of his decisive contribution. His 'rules by which to judge of causes and effects'—along with all of his other normative recommendations from *Treatise* 1.3—go completely unmentioned in *Treatise* 2.3.1–2, and likewise in the entire *Abstract*. Admittedly, in *Enquiry* 8 he does at certain points advocate systematic and normatively disciplined causal investigation (notably at *E* 8.13–15). But he never suggests that these norms are what require us to ascribe the same necessity to the physical and moral realms, or that Clarkean libertarians who insist on a distinction between physical and moral necessity are guilty of a breach of scientific good practice in failing to apply his rules consistently. Their fault is far more straightforward and decisive: either misunderstanding their own ideas, or making assertions that lack meaning for want of an appropriate idea. The argument clearly hinges, in other words, on the limits of thinkability, not on such things as scientific norms.

Nor can Beebee claim independent support from her preferred interpretation of Hume's two definitions, seeing them as 'descriptions of the two different *ways* in which we come to make . . . causal judgements' (p. 419). In her book she suggests a slightly different formulation, whereby the definitions 'exhaust the *reasonable* means by which we can come to make causal judgements' (*Hume on Causation*, p. 107), thus building normativity into the definitions themselves. But although much of her discussion here is illuminating and insightful, the textual foundation for such an interpretation is too weak to bear much interpretative weight.<sup>34</sup> Even if it is accepted, moreover, it does little to vindicate Hume's argument from *Def* and *CCI* to *Nec*. If we are to understand *Def*—'Necessity is to be defined in terms of constant conjunction and inference only'—as specifying *ways in which causal judgements are to be made*, rather than as *defining what necessity is*, then why should Hume's libertarian opponent accept the inference to *Nec*: 'It is not possible to maintain that there is a thick necessity in the operations of matter which is not present in the voluntary actions of intelligent beings'?



Samuel Clarke, for example, would fully accept that both physical events and human actions are causal, but would insist that the *kind* of causation involved is quite different, with absolute *physical* necessity in the one case, and mere *moral* necessity in the other.<sup>35</sup> It is not clear how a specification of *ways in which causal judgements are to be made*—even if agreed—can have any bearing on this question. If Hume's two definitions are just saying that our *basis* for causal judgements must take the form of either observed uniformities or natural inference, this seems completely silent on the question of whether or not there is a *single kind* of causation involved. It is only to the extent that causal judgements are required to involve a *single idea* that any such conclusion can potentially be drawn, and this, again, is clearly the main thrust of Hume's own argument.

It seems, then, that neither Beebee's appeal to the normativity of Hume's discussions of causation, nor her interpretation of the two definitions, can be of much assistance in making sense of his argument concerning liberty and necessity. Her reconciliation of that argument with New Humeanism, therefore, has to depend entirely on her initial guiding thought that for the sceptical realist, the very idea of necessary connexion—whose origin Hume traces—*itself* irresistibly represents a 'thick' power in the external objects themselves rather than the subjective impression from which it is copied. In her 2007 paper, Beebee appeals to this thought without explaining it at length, but her 2006 book (*Hume on Causation*, 176–8) gives a bit more detail, motivating it in what has become the standard way for New Humeans,<sup>36</sup> by comparison with Hume's treatment of the external world. Hume is generally considered to be a believer in 'the continu'd and distinct existence of body' (*T* 1.4.2.2), partly because he repeatedly says that we are naturally and irresistibly inclined to have such a belief (e.g., *T* 1.4.2.1, *E* 12.7–8), and partly because his other philosophical views—as expressed in his lengthy discussions of causal reasoning, morals, politics, economics, religion, and so on—seem to take it for granted. But if all our ideas of external objects are copied completely from subjective impressions, as implied by his Copy Principle (*T* 1.1.7, *E* 2.5), then it might seem impossible for us *even to form a thought* of a genuinely distinct and independent object. This suggests to Beebee that Hume's apparent acceptance of the external world must imply some loosening of his strict theory of ideas, some way in which 'our ideas and impressions [can] *represent* mind-independent reality, and thus represent that reality as a world of mind-independent chairs, tables, cats and dogs [even though] those ideas are *inadequate* ideas of what they represent'.<sup>37</sup> Carrying this suggestion over to the case of causation, perhaps the idea of necessary connexion—copied though it is from the internal impression of 'customary transition of the imagination' (*T* 1.3.13.3, *E* 7.28)—can somehow *represent* the mind-independent necessity that supposedly underlies inductive uniformities, and thus provide a vehicle for a Causal realist belief. That, at any rate, is the essence of Beebee's proposed New Humean account.<sup>38</sup>



The most fundamental problem with this account of Humean representation is that it is very poorly supported by the texts, which give no hint of how either an internal impression—or an idea copied from such an impression—might come to represent some supposed objective necessity that lies beyond our experience. Nor are the comparisons with the case of the external world helpful or convincing. Consider first the suggestion that Hume's *impression* of necessary connexion can somehow refer beyond itself to a 'thick' Causal power, with the copied idea thus inheriting this objective reference. Not only is there no trace of any such theory in Hume's text, but also, it would run quite contrary to his insistence—when arguing towards his account of our belief in body in terms of imaginative fictions—that impressions are transparently open to view and thus unable to point to anything beyond themselves:

That our senses offer not their impressions as the images of something *distinct*, or *independent*, and *external*, is evident; because they convey to us nothing but a single perception, and never given us the least intimation of any thing beyond. . . . all sensations are felt by the mind, such as they really are . . . every impression, external and internal, passions, affections, sensations . . . appear, all of them, in their true colours, as impressions or perceptions. . . . For since all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear. Every thing that enters the mind, being in *reality* a perception, 'tis impossible any thing shou'd to *feeling* appear different. (*T* 1.4.2.4–7)

In *Treatise* 1.3.14, Hume draws the obvious moral, that our impressions and ideas cannot possibly represent an objective 'thick' necessity, power, or efficacy, since we never perceive any such thing:

Ideas always represent their objects or impressions; and *vice versa*, there are some objects necessary to give rise to every idea. If we pretend, therefore, to have any just idea of this efficacy, we must produce some instance, wherein the efficacy is plainly discoverable to the mind, and its operations obvious to our consciousness or sensation. (*T* 1.3.14.6, cf. 1.3.14.11)

There is, then, nothing new either discover'd or produc'd in any objects by their constant conjunction, and by the uninterrupted resemblance of their relations of succession and contiguity. But 'tis from this resemblance, that the ideas of necessity, of power, and of efficacy, are deriv'd. These ideas, therefore, represent not any thing, that does or can belong to the objects, which are constantly conjoin'd. This is an argument, which . . . will be found perfectly unanswerable. (*T* 1.3.14.19)



Even if we abandon this Humean doctrine of transparent representation, and adopt some more sophisticated causal theory of reference for impressions (whereby an impression can represent the hidden underlying reality from which it arises), it is still not clear that Beebee's theory can be adequate to the task of explicating Hume's argument concerning liberty and necessity. For that argument is supposed to *guarantee* that our impression and idea of necessity refer to a *single kind* of necessity, and this sort of causal theory could give no such guarantee. According to Hume, the impression arises from repeated observation—and consequent customary inference—both of material interactions (such as colliding billiard balls) and of human actions. His libertarian opponent, such as Clarke, claims that the *physical* and *moral* cases involve different kinds of necessity, while Hume aims to refute this by appeal to the impression which is common to both. But if all necessity is beyond our grasp except insofar as our thought represents it—inadequately—through the impression to which it gives rise, then no such refutation can work. For even if the libertarian agrees that exactly the same subjective impression is generated in both cases (through similar experienced uniformity and inference), he has no good reason to concede that this provides significant evidence—let alone decisive proof—of the same underlying necessity. Clarkean libertarians may be perfectly willing to accept that both physical and moral necessities give rise to uniformity (and thus inductive prediction). But unless that common manifestation constrains human thought—as it does on the Old Humean picture—it poses no threat to our supposing that there are different necessities at work in the physical and moral cases.

If an internal impression cannot represent an objective necessity, then an idea copied from such an impression fares no better. First, Hume repeatedly insists that our ideas 'exactly represent' the impressions from which they are copied (*T* 1.1.1.7, cf. 1.1.1.12, 1.3.7.5, 1.3.14.6, 1.3.14.11, *A* 6),<sup>39</sup> and can never represent anything else 'without a fiction' (*T* 1.2.3.11). Secondly, in the case of the external world, *but not in the case of necessity*, Hume gives an elaborate explanation of how such a 'fiction'—a sort of pseudo-idea—can arise without being put together from *bona fide* impression-copied ideas (*T* 1.4.2.36–43). Thirdly, Hume says that those who attempt to get beyond the confused fictions of the vulgar, by giving a philosophically respectable account of the external world, fail hopelessly to achieve this:

Philosophers deny our resembling perceptions to be identically the same, and uninterrupted; and yet have so great a propensity to believe them such, that they arbitrarily invent a new set of perceptions, to which they attribute these qualities. I say, a new set of perceptions: for we may well suppose in general, but 'tis impossible for us distinctly to conceive, objects to be in their nature any thing but exactly the same with perceptions. What then can we look for from this confusion of groundless and extraordinary opinions but error and falsehood? (*T* 1.4.2.56)



So it seems that even in the case of the external world, philosophers are after all limited in thought by the Copy Principle, trapped within the confines of their own perceptions.

All this seems a long way from Beebee's subtle New Humeanism, which purports to assign a content to our idea of necessity going well beyond the nature of the impression from which it is derived. And this contrast is particularly striking, given that Hume's discussion of that idea's origin—as we saw in §II—is explicitly undertaken with the aim of clarifying its significance by identifying its impression source:

We shall . . . endeavour, in this section, to fix, if possible, the precise meaning of these terms [*necessary connexion*, etc.] . . . [by] what invention can we throw light upon [our most simple] ideas, and render them altogether precise and determinate to our intellectual view? Produce the impressions or original sentiments, from which the ideas are copied. (*E* 7.3–4)

In respect of its appeal to the theory of ideas, therefore, Hume's treatment of necessary connexion is radically different from his treatment of the external world. So far from insisting that no corresponding impression can be found—and therefore having to resort to an imaginative 'fiction' to account for the problematic belief—here a *bona fide* impression is explicitly identified and trumpeted as providing the source and meaning of the idea of necessary connexion. Moreover, we are given no suggestion that anything beyond that impression need be sought in order to make philosophical sense of the idea. Admittedly some further content is present within the 'vulgar, inaccurate idea' of power, because this is typically contaminated with feelings of the 'animal nisus' or 'strong endeavour' that we experience when striving physically against resistance (*E* 7.15, n. 13). But such feelings 'can afford no accurate precise idea of power', and we are unambiguously guilty of an *error* if we 'transfer [them] to inanimate objects, and . . . suppose, that they have some such feelings' (*E* 7.29, n. 17).<sup>40</sup> Elsewhere, Hume hints at another possible source of contamination, in that the vulgar are inclined:

to imagine they perceive a connexion betwixt such objects as they have constantly found united together; and because custom has render'd it difficult to separate the ideas, they are apt to fancy such a separation to be in itself impossible and absurd. (*T* 1.4.3.9)

But again, this seems to be dismissed as an error, rather than being embraced as a potential source of illumination regarding the cognition—or even the mere thought—of genuine powers.<sup>41</sup>

In short, from the point when he identifies the crucial impression for which he has been seeking, Hume is unequivocal that all *legitimate* thought



about necessary connexion must answer to the corresponding idea and to the two definitions that characterise its circumstances of origin:

Necessity, then, . . . is nothing but an internal impression of the mind . . . Without considering it in this view, we can never arrive at the most distant notion of it, or be able to attribute it either to external or internal objects. . . . (T 1.3.14.20)

These two circumstances form the whole of that necessity, which we ascribe to matter. Beyond the constant *conjunction* of similar objects, and the consequent *inference* from one to the other, we have no notion of any necessity or connexion. (E 8.5)

Such passages state clearly that only ‘thin’ necessity can be *attributed* or *ascribed*. Yet on Beebee’s New Humean theory, our idea of necessary connexion is supposed to represent a ‘thick’ necessity, and we are even enjoined to apply that idea reflectively—to cases where the impression does not arise naturally—on the basis of Hume’s ‘rules by which to judge of causes and effects’ and so on. Such selective application would seem precisely to involve the *ascription* of thick necessity, in direct contradiction of Hume’s words. Beebee attempts to elude this sort of objection by drawing a distinction between *ascription* and *reference* (p. 422), but such a distinction seems very unconvincing. There is surely a serious oddity in the suggestion that we can *refer* to and *selectively apply* something of which we have ‘no notion’ (indeed, not even ‘the most distant notion’), and which we consequently cannot *attribute* or *ascribe*. Absolutely nothing in Hume’s texts justifies the claim that he would distinguish between the different kinds of representation in this way.<sup>42</sup> Overall, therefore, it is hard to see how Hume’s theory of ideas could possibly accommodate what Beebee’s New Humeanism requires, namely, that the idea of necessary connexion—though admittedly derived from a subjective impression—can nevertheless represent a ‘thick’ necessity that goes beyond his two definitions.

The failure of Beebee’s attempt to reconcile Causal realism with Hume’s argument concerning liberty and necessity has a wider moral. Her aim was to exhibit a form of ‘sceptical realism’ which could achieve such a reconciliation, and she accordingly adopted an approach which—like Hume’s own argument—gave a central role to the idea of necessary connexion that he identifies. Given this background, Beebee went on to argue that ‘Hume’s discussion of free will provides virtually no *additional* evidence’ (p. 413, my emphasis) against such Causal realism, and to some extent I agree. *If* it were plausible to read Hume as saying *both* that our understanding of necessary connexion is completely captured by that idea, *and also* that the idea determinately represents an objective ‘thick’ necessity, then her reconciliation might come close to success.<sup>43</sup> But no such plausible reading is available: there is nothing in Hume’s texts to suggest that the subjective idea can



indeed represent a ‘thick’ necessity in the way that she requires. Hence the decisive objections to her reconciliation are to the form of ‘sceptical realism’ on which it is based, rather than specifically to her reading of Hume’s argument concerning liberty and necessity. As we shall see, this contrasts with the position with regard to Peter Kail and John Wright, who favour an interpretation of ‘sceptical realism’ in which the idea of necessary connexion plays a far less central role, but who accordingly encounter even greater difficulties in making sense of Hume’s discussion of liberty and necessity.

## V. KAIL ON LIBERTY AND NECESSITY

Peter Kail’s response to the challenge posed by ‘Of Liberty and Necessity’ is quite different from Beebee’s; indeed, the first half of his paper on the topic is devoted to an analysis of Hume’s ‘sceptical realism’ which is in stark contrast to her ‘natural belief’ account. According to Kail:

Realists [i.e., New Humean interpreters] do not, and more importantly, need not, read Hume as *believing* in powers.<sup>44</sup> Realists instead talk of Hume’s ‘assumption’ or his ‘taking for granted’ or his ‘supposition’ of powers. . . . And realists typically do not say that the attitude towards the existence of power admits of what, straightforwardly at least, is *epistemic justification* . . . Realism need not involve belief or epistemic justification: so how then are we to understand it?<sup>45</sup>

He answers his own question by insisting that this ‘realism’ can only be understood by contrast with what it opposes:

At the forefront of the original realist readings of Strawson, Edward Craig and John Wright is an anti-realism fuelled by an understanding of Hume’s story about impressions and ideas . . . the focus of which is conceptual or semantic. . . . At a minimum, realism holds that we can form thoughts that reach beyond the deliverances of impressions and thereby allow for the possibility of an ontology that includes genuine causal power or external objects. . . . (pp. 254–5)

Thus, as we saw in §I, Kail’s minimal requirement for ‘sceptical realism’ is compatible with complete agnosticism about ‘genuine causal power’, and a long way from Beebee’s interpretation with its unshakable, imagination-compelled natural beliefs.

Kail’s treatment also contrasts strongly with Beebee’s in putting far less emphasis on the impression and idea of necessary connexion that Hume himself identifies, and Kail would apparently dismiss her suggestion that this subjective idea might somehow represent an objective power:



... as all realists acknowledge, we have no *idea* of necessity derived from a genuine experience of power. Hume's official account of the state of belief implies that any such belief in causal power would have to involve such an idea. Therefore no such belief is available to Hume. (p. 254)

On Kail's interpretation, the kind of 'thought ... relevant to Hume's realism about causation' does not involve 'an *idea* of necessity or [even] a relative idea of necessity'. Instead, it 'expresses itself in how Hume understands what it would be to have a *genuine* impression of power':

If we were to perceive power—have an impression of it—we would be (a) able to 'read off' what effect some object must have and (b) find it impossible to conceive of the cause without the effect. So when asked what is one thinking of when one thinks of power, the appropriate answer is that which, were we to grasp it, would furnish the capacity for such '*a priori*' inference and close down our powers of conception. Call this the reference-fixer for 'power' (RFP). ... We have no understanding of what feature it is that would yield those consequences. ... It is a thought of a kind that manifests itself in Hume's argumentative strategy. (p. 256)

Having explained his RFP notion, Kail acknowledges that it is in some tension with familiar elements of the Humean package:

... one might find [appeal to the RFP] objectionable in a number of ways. First one might argue that the alleged feature is, so specified, incoherent: it is impossible for there to be any such feature since we can *always* separately conceive cause without effect and we can *never* infer *a priori* effect from cause. ... Second, one might argue that the grasp of the RFP manifested in Hume's arguments is merely for *reductio*, and so signals no endorsement of it. ...

I have argued elsewhere that both of these anticipated objections are indeed very serious. First, Kail's RFP notion comes badly into conflict with Hume's oft-repeated 'Conceivability Principle', a principle whose application Kail attempts to limit to 'sensory experiences or *impressions*'<sup>46</sup> but on a very thin textual basis.<sup>47</sup> Secondly, there is little solid evidence for viewing Kail's RFP as carrying implicit endorsement within 'Hume's argumentative strategy', which can be explained perfectly well in other ways.<sup>48</sup> But since my aim in this chapter is to focus on the specific difficulties for New Humeanism arising from Hume's treatment of liberty and necessity, let us put aside these worries here.

Kail goes on to address the question of how realist thought as characterised by his RFP can arise experientially in accordance with Humean



principles, and suggests that this can happen because customary inference is able to ‘mimic a genuine experience of power’:

First, the determination of the mind effects an *immediate* and *non-reasoned* transition from cause to effect . . . But that immediacy is the effect of prior habit and not of a grasp of causal power. Second, the incapacity to conceive cause without effect that a true grasp of necessity involves is given a psychological twist by Hume. The effect of the repeated experience of A and B is finding it psychologically impossible to think of the cause object without its effect. . . . (p. 258)

Recall that we saw serious problems in Beebee’s account, in that it left quite mysterious how Hume’s subjective ‘idea of necessary connexion’ could possibly give content to a belief in an objective thick power. Kail here proposes a clever way of trying to fill this gap, backed up by a passage that we have already encountered in §IV:

’Tis natural for men, in their common and careless way of thinking, to imagine they perceive a connexion betwixt such objects as they have constantly found united together; and because custom has render’d it difficult to separate the ideas, they are apt to fancy such a separation to be in itself impossible and absurd. (*T* 1.4.3.9).

After quoting this passage, Kail continues:

Fancying the ‘separation’—the independent conception of the cause without the effect—is something we come to think of as *impossible*: which is precisely what the RFP assumed. That gives us a psychological explanation of immediate transition and the impossibility of conceiving cause without effect without our ever having had a genuine experience of power. . . . But it is important to note that our actual *idea* of necessity—the determination of the mind—does not itself *represent* power in the objects (and neither does the impression). For that idea is a copy of the relevant impression, which is simply an impression of our *psychological* determination. . . . [However] a grasp of the RFP itself can emerge from this idea of necessity: the view that power would involve immediate inference and the incapacity to conceive cause without effect. This notion of power is ‘fancied’ in virtue of the kind of effect the customary transition has upon our minds, though the RFP is not *identical* to that idea. It is something we can suppose of causal powers even when we have no idea of *what* would yield the cognitive consequences so specified nor the faintest notion of *what* power might be. (p. 259)

This kind of account—which aims to explain how the imagined inseparability of ideas in causal inference can yield an illusion of grasping an a



priori connexion—is attractive to New Humeans, because it neatly links Hume’s account of causal reasoning with the absolute necessity that they take to characterise ‘real’ powers.<sup>49</sup> But claiming it as a core component of Hume’s own philosophy of causation by appeal to *T* 1.4.3.9 seems very unconvincing, because there he is explicitly criticising the ‘antient philosophers’ for following a vulgar error of the imagination in ascribing objective powers to objects: in other words, he is giving an *error theory*, not an account of a notion that he sees as potentially legitimate. Later in the same paragraph he says that the ‘just inference’ would be ‘that we have no idea of power or agency, separate from the mind, and belonging to causes’. Moreover, trying to identify such power or agency in objects is to ‘seek for it in a place, where ’tis impossible it can ever exist’. He goes on in this vein for the rest of the section, and indeed a natural reading of *T* 1.4.3.9–10 would put the New Humeans’ supposition of ‘genuine power’ in the same boat as the antient philosophers’ ‘faculties’ and ‘occult qualities’, for which Hume shows nothing but contempt.

In his book (*Projection and Realism in Hume’s Philosophy*, 107), Kail cites other passages that mention, or at least hint at, something like the feeling of inseparability that is described most explicitly at *T* 1.4.3.9. Thus at *T* 1.3.6.15, Hume says that ‘we have no other notion of cause and effect, but that of certain objects, . . . which in all past instances have been found inseparable’. At *T* 1.3.8.13, he speaks of custom operating on the mind ‘before we have time for reflection. The [cause and effect] seem so inseparable, that we interpose not a moment’s delay in passing from the one to the other’. In the *Enquiry*, he suggests that after someone ‘has observed several instances’ of the conjunction of cause and effect, ‘he now *feels* these events to be *connected* in his imagination, and can readily foretel the existence of one from the appearance of the other’ (*E* 7.28).<sup>50</sup> But again, none of these passages does much to support the claim that Hume sees such feelings of inseparability and immediacy of inference as giving a handle on a genuine and objective ‘thick’ notion of causation; rather, the passages in context seem mainly to be providing an insightful description of an admittedly misleading aspect of human phenomenology.

Quite apart from this weak textual basis, there is also a particular philosophical difficulty in Kail’s suggestion that Hume might envisage a distinct (and respectably truth-apt) notion of necessary connexion as arising through our grasp of his RFP: as *something* ‘*which, were we to grasp it, would furnish the capacity for such a priori inference and close down our powers of conception*’. For by weakening Hume’s Copy-Principle empiricism, this method of extending our conceptual repertoire would appear to undermine his own objection to Locke’s attempt to explain the origin of the idea of power as *something capable of producing changes in material things* (*T* 1.3.14.5 and *E* 7.8, n. 12, referring to Locke’s *Essay* II xxi 1).<sup>51</sup> If, even in the absence of a relevant impression, we can form a respectable notion of something which would yield certain *cognitive* consequences, then it is



hard to see why we could not form an equally respectable notion of something which would yield certain *material* consequences. Of course this is not by itself decisive against Kail's account—it could be that Hume just failed to note the inconsistency—but it is certainly a substantial objection.

Having given his explanation of how Hume can admit an RFP-based notion of necessity, Kail suggests that Hume indeed endorses that notion, on the basis of 'a crucial passage' (i.e., *T* 1.4.4.1) which he interprets as stating that 'the "customary transition from causes to effects" that is the idea of *necessary connection*' is a principle of the imagination which is 'permanent, irresistible and universal' (Kail, p. 261). 'Any authority that that inferential disposition has', Kail continues, 'can lend minimal but nevertheless genuine support to an assumption that there is power underlying the regularities. This minimal support, and Hume's preparedness to talk of hidden connections, tips the balance in favour of realism.' Thus Hume 'shows a preference in favour of a metaphysic of power', but Kail is very careful to insist on its minimal nature and scope:

This preference concerns solely what occurs at the level of metaphysics and plays no role in common life. For those purposes, the existence of power or otherwise is irrelevant. It is partly for this reason that such an 'assumption' or 'supposition' does not manifest itself in the psychological state of *belief*. Beliefs tend to govern one's behaviour in a way that assumptions do not. . . . The minimal preference for realism is simply a preference for what metaphysical position the combination of our natural propensities and reason delivers: reason suggests agnosticism but that is trumped by the natural propensities. (pp. 261–2)

Again, one can raise problems; for example, Kail's 'crucial passage' at *T* 1.4.4.1 does not even mention necessary connexion, but focuses on custom, whose central role in Hume's philosophy is already uncontroversial (so the passage adds nothing of relevance here). Customary inference is indeed sometimes identified by Hume with an *impression* of necessity (never an *idea*, as Kail suggests), but it is anyway subjective, so it can only dubiously 'lend support to an assumption' of objective power. Moreover, Section 1.4.4 of the *Treatise* finishes by concluding that the two 'permanent, irresistible and universal' principles identified in its first 'crucial' paragraph are in 'direct and total opposition', so it is somewhat problematic to see that paragraph as providing the reflective support that Kail seeks.

We have by now encountered a number of formidable problems for Kail's general account of Hume's Causal realism, but let us here set them all aside, and move on to examine his approach to the specific difficulties of accommodating Hume's argument on liberty and necessity within a New Humean perspective. Although his sketch of the resulting objections is rather different from my own in §III, we can largely ignore the differences here. His reply starts as follows:



These objections (a) rest on the assumption that Hume's realism—the minimal preference for a metaphysics of genuine necessitation—must shape, or figure in, his reconciliation project [i.e. Section 8 of the *Enquiry*] and (b) crucially miss the fact that Hume refigures the dispute at the level of *common life* rather than as an issue in the metaphysics of causation. What this means is that Hume's realism—the minimal preference—is irrelevant to the discussion and so cannot be *inconsistent* with it. (pp. 263–4)

These two points together would indicate that we should not expect Hume's metaphysics of causation and necessity to figure in his argument concerning liberty and necessity, thus enabling the New Humean account to escape any responsibility for accounting for the logic of that argument. But the points themselves clearly require justification, and in fact point (b) in particular is very hard to square with Hume's texts. Kail introduces it by stating that the 'opening paragraph' of *Enquiry* 8 'implies a distinction between unsolvable metaphysical disputes and those of common life, and what is interesting is [that] the dispute between the libertarian and the necessitarian [is] deemed to be a debate in common life' (p. 264). Such a reading, however, goes well beyond the text. In the first paragraph of Section 8, Hume is clearly aiming to motivate his project of resolving the free-will issue through clarification of the relevant ideas, by emphasising the plausibility of his claim that the issue has persisted because of misunderstanding and ambiguity. One might naturally expect, he suggests, that in any very long-running debate the meaning of the terms would have been agreed upon. But on the contrary, the very fact that a debate has remained undecided for so long indicates that there is likely to be some ambiguity in the terms involved. Of course this might not always be the case:

It is true; if men attempt the discussion of questions, which lie entirely beyond the reach of human capacity, such as those concerning the origin of worlds, or the oeconomy of the intellectual system or region of spirits, they may long beat the air in their fruitless contests, and never arrive at any determinate conclusion. But if the question regard any subject of common life and experience; nothing, one would think, could preserve the dispute so long undecided, but some ambiguous expressions. . . .

He then starts the next paragraph by saying that 'This has been the case in the long disputed question concerning liberty and necessity', in which 'a few intelligible definitions would immediately have put an end to the whole controversy' (*E* 8.2).

Hume thus confirms that the question of liberty and necessity 'regards common life and experience', and implicitly denies that it lies 'entirely beyond the reach of human capacity' (as indeed he intends to demonstrate).



But nothing here suggests a general dichotomy ‘between . . . metaphysical disputes and those of common life’. Indeed, the word ‘metaphysics’ and its cognates do not even occur here,<sup>52</sup> and Hume’s usage elsewhere in the *Enquiry* (e.g., *E* 1.7, 1.12, 7.2–3) indicates that he anyway does not take metaphysical disputes to be typically unsolvable. Nor, even, do his two uses of the word within Section 8 give any support to Kail’s reading. At 8.23 Hume refers to ‘the question of liberty and necessity; the most contentious question . . . of metaphysics’, and at 8.27 he pairs ‘metaphysics’ with ‘natural philosophy’ against ‘morality and religion’ in respect of the kind of necessity traditionally understood to apply there. Thus we have an explicit statement that the question of liberty and necessity is indeed metaphysical (*contra* Kail), and an implicit indication that this very metaphysical question stands to be illuminated by clarification of the relevant terms in much the same way as questions of natural philosophy. There are also other strong indications within the section against Kail’s suggestion that Hume is operating with a general distinction between common life and metaphysics. Thus at 8.16 he points out that the ubiquity of constant conjunction has never been disputed ‘either in philosophy or common life’, and at 8.27 he makes a similar claim about constant conjunction and inference ‘in the schools, in the pulpit, and in common life’. These suggest a *continuity* between philosophy and common life, rather than any sharp dichotomy. There is only one other mention of ‘common life’ in the section, in the very final paragraph, where Hume contrasts ‘the sublime mysteries’ of divine foreknowledge and theodicy with ‘the examination of common life’. This certainly fits with the contrast drawn in the first paragraph, but again does nothing to suggest that ‘common life’ is to be contrasted with *metaphysics* in general.

There is also the obvious point that Hume’s argument concerning liberty and necessity is presented in three texts, while his introductory comment on disputes in common life is entirely absent from the *Treatise* and *Abstract*. Kail confronts this objection in a puzzling footnote:

Objection: this [the mention of common life] only occurs in the first *Enquiry*. Since similar problems for Hume’s realism can be mounted on the basis of Hume’s *Treatise* discussion of liberty and necessity, appeal to this passage to remove the objection cannot help the problem formulated in the *Treatise*. Response: all this means is that the *Enquiry* affords a better case for realism. Realist readers (e.g. Strawson 1989: 8) view the *Enquiry* as superior to the *Treatise* with respect to the discussion of causation in this respect not least because the references to secret powers are more prominent, so such a move is not *ad hoc*. Those who prefer the first *Enquiry* to the *Treatise* thus have reason to take this as authoritative. (p. 268, n. 26)

Kail’s reply here misses the fundamental point, that Hume’s discussions in the *Treatise* and *Abstract* appear to present *essentially the same argument*



as in the *Enquiry*.<sup>53</sup> So if the appeal to common life is ‘crucial’ to a correct understanding of that argument, as Kail claims, then this leaves the *Treatise* and *Abstract* as crucially deficient. But to judge them as such seems *prima facie* implausible: nothing within them gives any indication that they are so deficient. Hence we have strong reason to prefer an interpretation of the *Treatise* and *Abstract* discussions that does not rely on Kail’s ‘common life’ claim, and also strong reason—given the apparent identity of the argument—to do the same for the *Enquiry*. Kail’s footnote does nothing to answer this straightforward point.

Leaving these criticisms aside, let us finally see how Kail deals with the fundamental objection of §III above: that the strategy of Hume’s argument for ascribing exactly the same necessity to matter and mind seems to turn on the denial of any kind of thick necessity that goes beyond his two definitions. Kail appreciates the force of this objection, acknowledging that it ‘would be devastating—if it weren’t for the fact that there are clues in [*Enquiry* 8] that this cannot be Hume’s strategy since he himself violates it’. (p. 266).<sup>54</sup> Kail describes the alleged violation as follows:

But now to the main point: in the midst of the discussion Hume is prepared to grant, for the sake of argument, power in matter, but that it makes no difference to the reconciliation. But if the reconciliation turned on the claim that no further thought is possible with regard to causation, even this small concession would violate this alleged central move. The context here is where Hume considers those who ‘maintain that it is possible to discover something farther in the operations of matter’ [*E* 8.27] in opposition to the mild positivism of Hume’s two ‘definitions’. Here is an opportunity for Hume to reassert his alleged conclusion that no such thing is [possible] since no genuine thought is possible and use his ‘theory of meaning’ to fix the metaphysics. But he does not take this opportunity: instead he tells us that such a concession is ‘of no consequence to morality or religion, whatever it may be to natural philosophy or metaphysics’. (p. 266)

There are at least three objections to be made here. First, as should be clear from the outline in §III, *Enquiry* 8.27 cannot reasonably be described as ‘in the midst of the discussion’, at least if this is taken to imply that it occurs within Hume’s main argument. That main argument runs from 8.4 to 8.20, corresponding closely to the content of Section 2.3.1 of the *Treatise*.<sup>55</sup> By the time we reach 8.27, Hume is focusing very specifically on the religious and moral implications of the doctrine of necessity, and it is within this context that he ‘is prepared to grant, for the sake of argument, [thick] power in matter’. The second objection to Kail’s gloss on the paragraph is that Hume does *not* say here that such a concession of thick power in matter ‘makes no difference’ to the ‘reconciling project’ (*E* 8.23) of Section 8. He only insists that it ‘can be of no consequence to morality or



religion, whatever it may be to natural philosophy or metaphysics', a wording which clearly suggests—*contra* Kail—that the concession *would indeed* be of some consequence to natural philosophy and metaphysics. Thirdly, Kail is mistaken in saying that Hume 'does not take this opportunity . . . to reassert his . . . conclusion that no such . . . thought [of a thick necessity in matter] is possible'. For immediately after the passage that Kail quotes, Hume's very next sentence begins 'We may here be mistaken in asserting, that there is no idea of any other necessity or connexion in the actions of body.' Here he is precisely reasserting his conclusion about the limits of our thought, and reminding his reader that the discussion of the present paragraph (concerning the moral and religious consequences of his position), though it does not rest on that conclusion, is entirely without prejudice to it.

Kail's gloss on *E* 8.27 is thus seriously at odds with Hume's words: a careful reading tells against his interpretation rather than for it, and in more than one way. That aside, he has clearly failed to achieve what he hoped in citing it, namely, to identify a violation of the argument strategy outlined in §III above. In the absence of any such violation, we have every reason to interpret Hume as following this clear and straightforward strategy, which seems to fit excellently with his text in the *Treatise*, the *Abstract*, and the *Enquiry*. So the fundamental objection to the New Humean interpretation based on 'Of Liberty and Necessity', which Kail himself acknowledges to be potentially 'devastating', retains its full potency—it was, and so far remains, devastating.

## VI. WRIGHT ON LIBERTY AND NECESSITY

John Wright's interpretation of Hume on liberty and necessity appears in Chapter 5 of his 2009 book, *Hume's 'A Treatise of Human Nature'*, building on his Chapter 3, which outlines a position on causation that is in many ways similar to Kail's. On this account, 'the natural supposition of objective necessity' (pp. 122–6) is not a standard Humean belief (i.e., an enlivened idea), but is more like our belief in a necessary truth in which we are 'determin'd to conceive [the ideas] in [a] particular manner' (*T* 1.3.7.3)—it is this felt determination, rather than the vivacity of any idea, that constitutes our assent or belief in causal necessity. A 'natural supposition of inseparability' (pp. 124–6) arises when 'custom and association make us suppose that the cause and effect are conceptually inseparable, just as they would be if we had *insight* into their real natures' (pp. 124–5). Wright thinks that this explanation of our belief in objective necessity, for which he has recently coined the term 'quasi-inseparability',<sup>56</sup> dominates Hume's account of necessity in the *Enquiry*, but coexists in the *Treatise* with an alternative explanation based on projective 'spreading of the mind' (pp. 122–4; cf. note 40 of the current paper).



In making the case for this quasi-inseparability interpretation, Wright—like Kail—cites the *Treatise* passage in which Hume says that people have a natural tendency ‘to imagine they perceive a connexion betwixt’ cause and effect, ‘and because custom has render’d it difficult to separate the ideas, they are apt to fancy such a separation to be in itself impossible and absurd’ (*T* 1.4.3.9). He also cites some additional passages from the *Treatise*,<sup>57</sup> but puts particular weight on *Enquiry* 7:

As Hume puts it in his first *Enquiry*, after experiencing a constant conjunction between two successive events we ‘suppose, that there is some connexion between them; some power in the one, by which it infallibly produces the other, and operates with the . . . strongest necessity’ [*E* 7.27]. . . . This interpretation was further developed by Norman Kemp Smith, who argued that, according to Hume, we apprehend the objective necessity between objects through *feeling* rather than reason. This is the natural way to read Hume’s claims about the feeling which is the origin of the idea of necessity in ‘Of liberty and necessity’ in Book 2 of the *Treatise*: there he writes that while it is ‘impossible for the mind to penetrate’ into the relation of the objects, after experiencing their constant conjunction, the mind ‘*feels* the necessity’ [*T* 2.3.1.16]. Even more clearly, in his first *Enquiry* he writes that a person ‘*feels* . . . events to be *connected* in his imagination’ after experiencing their constant conjunction [*E* 7.28]. (pp. 125–6)

Hume’s painstaking investigation ‘Of the Idea of Necessary Connexion’ indeed seems the right place to look for convincing support. By contrast, passages scattered around the *Treatise* that happen to mention something like imagined or felt inseparability can give only very weak evidence of Hume’s view, unless they are appropriately linked to his own account (which *T* 1.4.3.9 is not—cf. §V). But I am not persuaded that the *Enquiry* 7 passages Wright quotes have the significance he gives them. In context, when Hume says that someone ‘*feels* . . . events to be *connected* in his imagination’ (*E* 7.28), he seems to be referring to exactly the same ‘connexion, . . . which we *feel* in the mind, this customary transition of the imagination’ which he has identified earlier in the same paragraph as ‘the sentiment or impression, from which we form the idea of power or necessary connexion’. Likewise, when he says that the mind ‘*feels* the necessity’ (*T* 2.3.1.16), in context he does not appear to be saying that ‘we apprehend the objective necessity between objects through *feeling*’ (Wright’s gloss on Kemp Smith). On the contrary, that entire paragraph seems to stress that the necessary connexion is not *apprehended* objectively at all, but ‘is merely a perception of the mind’. Of course these doubts by themselves do not totally undermine Wright’s account, which must be assessed on the basis of how well it can make sense of Hume’s treatment of causation and necessity in general,



and that is too big a task to attempt here.<sup>58</sup> Rather, my aim in this chapter is to focus on the significance of Hume's discussions of liberty and necessity in particular, so let us move on immediately to what Wright has to say in his Chapter 5.

Wright—whose 2009 book focuses on Hume's *Treatise*—acknowledges from the start that the account of causation (*T* 1.3.14) plays a crucial role in the treatment of liberty and necessity (*T* 2.3.1–2), and he sets out his stall accordingly:

A central question which arises here . . . is whether Hume reduces the meaning of causal necessity to uniformity and predictability—or whether he regards the latter as a sign of a genuine but unintelligible necessity. I shall argue that the latter assumption runs throughout his reasoning . . . (p. 170)

'Hume begins his discussion' in *Treatise* 2.3.1, says Wright (p. 177), 'by considering the necessity which we ascribe in the case of the physical world':

'Tis universally acknowledg'd, that the operations of external bodies are necessary, and [without] the least traces of indifference . . . Every object is determin'd by an absolute fate to a certain degree and direction of its motion, and can no more depart from that precise line, in which it moves, than it can convert itself into an angel . . . (*T* 2.3.1.3; cf. *E* 8.4)

Wright asks us to note that in this passage, 'Hume unambiguously ascribes necessity to external objects themselves' (rather than to 'the mind of the spectator'), and also 'describes this necessity as absolute and as excluding the possibility of anything occurring besides what actually occurs in the given circumstances . . . We appear to have necessity ascribed to objects in the strongest sense of the word'.<sup>59</sup> This ascription 'is not based on any penetration "into the essence . . . of bodies" . . .',

Rather the belief in the necessity of the operations of physical objects arises from the observation of their 'constant union' [which gives rise to] our psychological propensity to infer the one from the other . . . Hume sets out in the rest of this section to 'prove from experience, that our actions have a constant union with our motives, tempers, and circumstances', and that we infer that future actions follow necessarily from this regularity. (pp. 178–9)

Wright thus interprets the argument of *Treatise* 2.3.1 as being 'that we have exactly the same grounds to ascribe necessity to human actions as we have to ascribe it to material objects; since we ascribe it to material objects, we must also ascribe it to human actions' (p. 177). But this gives an *epistemic*



reading to the argument of the section, based on empirical *grounds*, and thus exactly of the kind that we earlier concurred with Beebee in rejecting.<sup>60</sup> A passage from *T* 2.3.1.4 would convey a somewhat different impression:

If objects had not an uniform and regular conjunction with each other, we shou'd never arrive at any idea of cause and effect; and even after all, the necessity, which enters into that idea, is nothing but a determination of the mind to pass from one object to its usual attendant, and infer the existence of one from that of the other. Here then are two particulars, which we are to consider as essential to necessity, *viz.* the constant *union* and the *inference* of the mind; and wherever we discover these we must acknowledge a necessity.

Admittedly, Hume is here discussing not only the origin and nature of our idea of necessity, but also the circumstances in which we ascribe it, thus introducing an epistemic element. It is only later, in *Treatise* 2.3.2, that the semantic theme clearly dominates, creating difficulties for Wright's reading.

The crucial section of Hume's argument is discussed by Wright under the heading 'The Theoretical Mistake of Those Who Believe in Free Will' (pp. 182–6). As he frankly acknowledges (p. 183), 'Hume appeals to the two definitions of cause which he gives in Book 1 . . . [which] tend to support the view that Hume reduces objective causation to mere regularity, and necessity itself to a subjective determination of the mind of the observer. Here in Book 2, Hume writes:'

I define necessity two ways, conformable to the two definitions of *cause*, of which it makes an essential part. I place it either in the constant union and conjunction of like objects, or in the inference of the mind from the one to the other. Now necessity, in both these senses, has universally, tho' tacitly, in the schools, in the pulpit, and in common life, been allow'd to belong to the will of man . . . (*T* 2.3.2.4)

After quoting further from this *Treatise* paragraph, Wright comments:

Here, Hume is explicitly defining 'necessity' itself, and not just 'cause' as he did in Book 1. Both definitions are said to be definitions of necessity. We get no qualifications about the definitions being 'foreign to the cause'. His aim here is clearly to disarm his opponents' arguments by getting them to admit that they accept necessity in the two senses he allows in his definitions.

Hume recognises that his opponents may 'refuse to call' what is allowed by his official definitions 'necessity', and may claim that when we ascribe necessity to matter we mean something stronger [*T* 2.3.2.4]. His answer is that even if there *is* such an unknown necessity in matter, it is not what he is ascribing when he claims that people's actions are



necessary. He is not claiming that our wills are moved by anything like the force which makes a billiard ball move when another one strikes it. . . . He is arguing, in effect, that he has taken the sting out of the tail of physical or material necessity, and applied this new domesticated breed of necessity to human actions. In accord with his two definitions, he writes, 'I place it in the constant union and conjunction of like objects, or in the inference of the mind from one to the other'. It is only in these two senses that he is ascribing necessity to both physical events and to human actions. (p. 184)

To his credit, Wright has put his finger on the key problem highlighted in §III, and on the key passage in the *Treatise* where this comes to the fore. But his immediate response is to question Hume's sincerity:

It is difficult to believe that Hume is being entirely ingenuous here, given his unequivocal statements in the first section of 'Of liberty and necessity' that in the physical world 'every object is determin'd . . . to a certain degree and direction of its motion' [*T* 2.3.1.3], and that there is no indifference or chance in the nature of things. Did he not begin his whole discussion by ascribing necessity to physical objects in a stronger sense than that authorized by his two definitions? Moreover, there is reason to question whether he really does believe that the force which moves our wills is of an entirely different kind than that which moves physical objects. . . . [In *T* 1.4.5] he suggests that mental causation is probably ultimately based on physical causation. (pp. 184–5)

The main point here—to which we return shortly—concerns Hume's commitment to determinism and his corresponding denial of indifference, but there is a subsidiary point involving physical and mental causation, where I think Wright is misinterpreting Hume's words. This misinterpretation first emerges in the preceding quotation from his page 184, where he says that Hume 'is not claiming that our wills are moved by anything like the force which makes a billiard ball move when another one strikes it'. This, I presume, is a gloss on the first sentence or two of the following passage:

Let no one, therefore, put an invidious construction on my words, by saying simply, that I assert the necessity of human actions, and place them on the same footing with the operations of senseless matter. I do not ascribe to the will that unintelligible necessity, which is suppos'd to lie in matter. But I ascribe to matter, that intelligible quality, call it necessity or not, which the most rigorous orthodoxy does or must allow to belong to the will. I change, therefore, nothing in the receiv'd systems, with regard to the will, but only with regard to material objects. (*T* 2.3.1.4)



In this passage, Hume is not at all relinquishing his claim (which Wright correctly judges to be his considered view) that the same necessity applies in both the physical and mental worlds. On the contrary, he is repeating it, by saying that the very same necessity which ‘the most rigorous orthodoxy . . . must allow to belong to the will’ (and no other) should *also* be ascribed to matter. He acknowledges that he is thus challenging ‘the receiv’d system . . . with regard to material objects’, by rejecting as ‘unintelligible’ the metaphysically heavyweight ‘necessity, which is suppos’d to lie in matter’. But he makes this point in passing, because his overt aim here—as announced in the previous paragraph—is to show that his view has no ‘dangerous consequences to religion and morality’. Thus understood, the entire passage makes perfect sense, expressing more or less exactly the same thoughts as the parallel passage of the *Enquiry* (E 8.27, quoted fully in §III above). So there is no basis here for interpreting Hume’s words as disingenuous.

Wright himself puts most weight on the issue of determinism and indifference, but his way of expressing it is tendentious:

The key philosophical question is whether Hume can legitimately claim that he is *only* ascribing necessity to objects in the two senses given by his two definitions. Does he not clearly deny any ‘liberty of indifference’ to both the physical and the mental world independently of the observer? . . .

. . . Consider again what Hume says about the unfortunate prisoner. He ‘discovers the *impossibility* of his escape, as well from the obstinacy of the gaoler, as from the walls and bars with which he is surrounded’ [T 2.3.1.17; Wright’s emphasis]. . . . Hume clearly uses the word ‘impossibility’ . . . here to refer to features of the objective world of agents and material bodies, and not mere features of the mind of the prisoner who observes them. . . . As the late John Yolton has argued, one cannot reduce the objective language of causality and necessity which Hume uses in his philosophical discussions of these topics to the language of regular successions of events. . . . (pp. 185–6)

Even if we agree with Wright on ‘the key *philosophical* question’, the key *interpretative* question is quite different. We might well share Yolton’s doubt that one can ‘reduce the objective language . . . of necessity . . . to the language of regular successions of events’. But then *we* are under little pressure to accept such a reduction, because we presumably *do not* share Hume’s apparently firm commitment to his simplistic Copy Principle empiricism.<sup>61</sup> Sometimes, it might be appropriate to infer that because a position is *philosophically* objectionable, it cannot be what Hume intends. But this is not such an occasion. He himself describes his subjectivist analysis of necessity as ‘the most violent . . . of all the paradoxes’ in the entire *Treatise*, which can only hope to ‘overcome the



inveterate prejudices of mankind . . . by dint of solid proof and reasoning' (*T* 1.3.14.24). He then goes on to express the vivid exclamations of incredulity that he anticipates in response (*T* 1.3.14.26). So this is clearly a case where we cannot rely on our own philosophical instincts to reveal what Hume is saying, but have to follow the logic of his 'solid proof and reasoning'. Wright finds it 'difficult to believe that Hume is being entirely ingenuous' in proposing his solution to the problem of liberty and necessity, based as it is on the claim that his two definitions give 'the very essence of necessity' (*T* 2.3.1.10, 2.3.2.2; cf. *E* 8.22, n. 18, 8.25, n. 19). Such difficulty is entirely understandable if one approaches the text (as I presume Wright does) with a firm commitment to thick necessity: denying its existence or even meaningfulness can indeed seem incredible. But this counts for little, given both Hume's open acknowledgement that he expects his view to appear wildly paradoxical, and also how well—on Old Humean principles—his argument on liberty and necessity meshes with the whole thrust of his discussion of the idea of necessary connexion (in all three works). None of this conflicts with his commitment to universal determinism or his denial of 'indifference',<sup>62</sup> and it seems extremely unlikely that Hume himself could have considered there to be any such inconsistency, when it is precisely his argument for determinism—the 'doctrine of necessity'—which turns so crucially on his two definitions.<sup>63</sup> Those who believe in thick necessity might indeed find some *appearance* of inconsistency, as for example in this passage cited by Wright:

'Tis universally acknowledg'd, that the operations of external bodies are necessary, and [without] the least traces of indifference . . . Every object is determin'd by an absolute fate to a certain degree and direction of its motion . . . (*T* 2.3.1.3)

Recall Wright's comment that here 'Hume describes this necessity as absolute and as excluding the possibility of anything occurring besides what actually occurs in the given circumstances . . . We appear to have necessity ascribed to objects in the strongest sense of the word' (p. 177). He thus suggests that this is a serious problem for the Old Hume interpretation, but it is not.<sup>64</sup> For if we take the Old Humean position seriously, then terms such as 'necessary' and 'determin'd', and even cognate terms such as 'impossibility', are to be interpreted through Hume's definitions. So if everything that happens is in accord with universal causal laws, then such things are indeed *necessary* as Hume understands that term, and contrary things are indeed (causally) *impossible*.<sup>65</sup> We can even agree with Wright that in the preceding passage, Hume intends to ascribe causal necessity 'in the strongest sense of the word', for this is precisely what Hume takes to be the strongest *legitimate* sense. New Humeans will of course insist against this that there is more to genuine, full-blooded necessity than Hume's two definitions allow, but we already know what his answer will be:



Beyond the constant *conjunction* of similar objects, and the consequent *inference* from one to the other, we have no notion of any necessity, or connexion. (E 8.5)

So although they may hanker after some more substantial notion of causal necessity, if (the Old) Hume is right, there is none to be had, and all of his declarations in favour of determinism and against ‘indifference’ must be interpreted accordingly. Perhaps this position is philosophically implausible, but it is where he was led by his Copy Principle empiricism, and for good or ill, it is clearly there in his texts, argued for systematically in the sections on the idea of necessary connexion, and correspondingly applied in the sections on liberty and necessity.

## VII. CONCLUSION

In the 2007 edition of Read and Richman’s collection *The New Hume Debate*, my chapter “Against the ‘New Hume’” ended with ‘a clear challenge’ to New Humeans to explain away the ‘apparently crucial semantic theme in Hume’s discussion of “liberty and necessity”’ (p. 247), suggesting that this would prove to be the decisive objection to their interpretation. Since I wrote this, there have been three attempts to answer the challenge, exhibiting very different approaches, but none in my view successful. Beebe takes the logic of Hume’s argument most seriously, attempting to expound a form of ‘sceptical realism’ which can accommodate it by treating Hume’s idea of necessary connexion as the means by which thick necessity is represented. Kail tries to limit the scope of the argument to ‘common life’, and also to identify nuances within it that would reveal Hume’s argumentative strategy to be very different from how it appears. Wright’s approach is most direct, clearly recognising the direction and force of Hume’s reasoning, but denying that he can sincerely mean it, on the basis that it would run counter to his necessitarian commitments.

Not only do all three approaches fail, as we have seen, but also, their variety tends to suggest that no other approach is likely to work either. If interpreted in the Old Humean manner, Hume’s argument concerning liberty and necessity is simple, logical, and sincerely intended: it has straightforward and explicit premises, and moves from them in a systematic way to an appropriately implied (and equally explicit) conclusion. It addresses precisely the issue that Hume has set himself to resolve, and proceeds in just the way he describes, through definition and the clarification of meanings, as he emphasises strongly both at the beginning and the end of his main argument:

... in the long disputed question concerning liberty and necessity  
... if I be not much mistaken, we shall find ... that a few intelligible



definitions would immediately have put an end to the whole controversy. (*E* 8.2-3; cf. *A* 34)

Here then is the advantage of definitions. Let any one *define* a cause, without comprehending, as a part of the definition, a *necessary connexion* with its effect; and let him shew distinctly the origin of the idea, expressed by the definition; and I shall readily give up the whole controversy. But if the foregoing explication of the matter be received, this must be absolutely impracticable. (*E* 8.25; cf. *T* 2.3.1.18)

On this Old interpretation, moreover, the argument has an obvious motivation within Hume's system, giving a clear purpose to his extended investigation of the idea of necessary connexion, and yielding a powerful conclusion with the potential both to refute Clarkean libertarianism and to authorise a thoroughgoing application of causal explanation within moral science—aims to which we know he was strongly committed.<sup>66</sup> By contrast, if one tries to interpret the argument in accordance with New Humean principles, then there seems to be no obvious 'best way' of doing so—either to match with the text, or with Hume's objectives—and it is up to the ingenuity of each interpreter to try to find ways of stretching or shrinking it to fit (by adding implicit claims, e.g., about representation or common life, or by discounting inconvenient passages). We thus find symptoms of what Imre Lakatos called a 'degenerating research programme', with a variety of more or less arbitrary interpretative epicycles being invoked to avoid outright refutation by the recalcitrant texts. Developments of this kind are not always unwelcome; indeed, such exploration of novel options can lead to real progress, and much of the work done within the New Humean paradigm—not least in the books by John Wright, Helen Beebe, and Peter Kail—has been genuinely illuminating.<sup>67</sup> The debate thus leaves a legacy of enduring value, but nevertheless these symptoms indicate that on the core question of Hume's metaphysics of causation, it has run its course. The New Hume has had a good innings, but now it is time to call it a day.<sup>68</sup>

## NOTES

1. Abbreviations to the works of David Hume are as follows. *T*: David Hume, *A Treatise of Human Nature: A Critical Edition*, vol. 1, edited by David Fate Norton and Mary J. Norton (Oxford: Clarendon Press, 2007, originally published 1739–40); *E*: David Hume, *An Enquiry Concerning Human Understanding*, edited by Peter Millican (Oxford: Oxford University Press, 2007, originally published 1748); *M*: David Hume, *An Enquiry Concerning the Principles of Morals*, edited by Tom L. Beauchamp (Oxford: Oxford University Press, 1998, originally published 1751); *HL*: David Hume, *The Letters of David Hume*, edited by J. Y. T. Greig, 2 vols. (Oxford: Clarendon Press, 1932).
2. This term became prominent in the debate through Blackburn's "Hume and Thick Connexions", *Philosophy and Phenomenological Research*, 50 (1990):



- 237–50; reprinted in *The New Hume Debate: Revised Edition*, edited by Rupert Read and Kenneth A. Richman (London: Routledge, 2000/2007), 100–12; revised as chapter 5 of Blackburn, *Essays in Quasi-Realism* (Oxford: Oxford University Press, 1993); reprinted in *Reading Hume on Human Understanding: Essays on the First Enquiry*, edited by Peter Millican (Oxford: Clarendon Press, 2002), 259–76.
3. Galen Strawson, *The Secret Connexion: Causation, Realism, and David Hume* (Oxford: Clarendon Press, 1989), 84–5.
4. E.g., *The Secret Connexion*, 13–15, 222–8.
5. Norman Kemp Smith, *The Philosophy of David Hume* (London: Macmillan, 1941), 91, 372–3, 387, 393, 401–2.
6. John P. Wright, *The Sceptical Realism of David Hume* (Manchester: Manchester University Press, 1983), 150–5.
7. *Sceptical Realism*, 126.
8. Kenneth P. Winkler, “The New Hume”, in Read and Richman, 52–74, originally published in *Philosophical Review*, 100 (1991): 541–79; Donald Livingston, *Hume’s Philosophy of Common Life* (Chicago: University of Chicago Press, 1984); Janet Broughton, “Hume’s Ideas about Necessary Connection”, *Hume Studies*, 13 (1987): 217–44; Edward Craig, *The Mind of God and the Works of Man* (Oxford: Clarendon Press, 1987); Michael J. Costa, “Hume and Causal Realism”, *Australasian Journal of Philosophy*, 67 (1989): 172–90; John Yolton, *Realism and Appearances* (Cambridge: Cambridge University Press, 2000); Stephen Buckle, *Hume’s Enlightenment Tract: The Unity and Purpose of An Enquiry Concerning Human Understanding* (Oxford: Clarendon Press, 2001); Peter Kail, “Projection and Necessity in Hume”, *European Journal of Philosophy*, 9 (2001): 24–54; “Is Hume a Causal Realist?”, *British Journal for the History of Philosophy*, 11 (2003): 509–20; *Projection and Realism in Hume’s Philosophy* (Oxford: Oxford University Press, 2007); “How to Understand Hume’s Realism”, in Read and Richman, 253–69.
9. “Is Hume a Causal Realist?”, 512.
10. “Is Hume a Causal Realist?”, 510, 513.
11. However, Kail’s way of presenting the issue could be considered tendentious: an Old Humean might be entirely happy to accept that Hume makes reference—in a sense—to ‘necessary connection in the objects’ and to ‘hidden powers’, but these would not be understood as referring to *thick* necessary connexions or powers. For more on this, see Peter Millican, “Against the ‘New Hume’”, in Read and Richman, 211–52, §3.5.
12. As elsewhere, e.g., Peter Millican, “Hume, Causal Realism, and Causal Science”, *Mind*, 118 (2009): 647–712, at 648, n. 4.
13. “The New Hume”, 53.
14. “Is Hume a Causal Realist?”, 512, n. 12; cf. “How to Understand Hume’s Realism”, 255.
15. “The New Hume”, 63–4.
16. In personal correspondence, Ken Winkler has suggested to me that ‘the clearest account of my views on this point comes in the long paragraph on p. 73 of the paper . . . where I compare Hume to Berkeley’.
17. For other relevant quotations (from *T* 1.3.14.4, 14, 27; *E* 7.26, 29), see Millican, “Against the ‘New Hume’”, §2.1, and “Hume, Causal Realism, and Causal Science”, §3.
18. For much fuller discussion of the New Hume debate in general, see Millican, “Against the ‘New Hume’”, and especially “Hume, Causal Realism, and Causal Science”.
19. The definitions are also applied in *Treatise* 1.4.5, ‘Of the Immateriality of the Soul’, though the interpretative lessons to be drawn there are less



- straightforward—see Millican, “Hume, Causal Realism, and Causal Science”, §7.
20. See, for example, Strawson, *The Secret Connexion*, 8, and “David Hume: Objects and Power”, in Read and Richman, 31–51, at 31–3; Buckle, *Hume’s Enlightenment Tract*, 194–5; Helen Beebee, *Hume on Causation* (London: Routledge, 2006), 221–5; Kail, “How to Understand Hume’s Realism”, 262, 268, n. 26; John P. Wright, *Hume’s ‘A Treatise of Human Nature’* (Cambridge: Cambridge University Press, 2009), 126.
  21. For the same account, expanded to include quotations from both works, see §8 of Millican, “Hume, Causal Realism, and Causal Science”, a sibling paper to the present one, likewise descended from a talk given to the April 2008 University of York Conference on Causation. Peter Millican, “Humes Old and New: Four Fashionable Falsehoods, and One Unfashionable Truth”, *Proceedings of the Aristotelian Society, Supplementary Volume 81* (2007), 163–99, §VIII presents the argument in a more structured form with relevant references.
  22. Here the close correspondence between the *Treatise* and *Enquiry* accounts breaks down, sometimes making it impossible to identify parallel passages—see Millican, “Hume, Causal Realism, and Causal Science”, 695–7, for a comparative account.
  23. Hume’s position is famously compatibilist in the *Enquiry*, aiming to reconcile ‘the doctrine of necessity’ with ‘the doctrine of liberty’ by demonstrating the only ‘reasonable sense, which can be put upon these terms; and that the whole controversy has hitherto turned merely upon words’ (*E* 8.3). In the *Treatise*, by contrast, he uses the term ‘liberty’ to mean *chance* or *indifference*, whose existence he denies (see *T* 2.3.1.3, 2.3.1.18, 2.3.2.1–2, 2.3.2.6–8).
  24. Given this dialectical context, there is no basis for taking Hume’s statement that he ‘may here be mistaken’ as expressing serious doubts, *contra* Yolton, *Realism and Appearances*, 129, 130.
  25. Kail, “Is Hume a Causal Realist?”, 510.
  26. Such neglect might perhaps be explicable in terms of the blinkers that tend to be imposed by philosophical fashion and the undergraduate curriculum (cf. Buckle, *Hume’s Enlightenment Tract*, 24–6). Free will is commonly treated as a topic in introductory metaphysics and moral philosophy, but relatively rarely in the history of philosophy. Meanwhile, most courses and general books on Hume tend to focus on Book I of the *Treatise*, neglecting both Book II and the *Enquiry*. But it still seems astonishing that specialist writers on *Hume on causation* should have given so little attention to ‘Of Liberty and Necessity’, when it so obviously contains the main application of his two definitions. As far as I know, the discussions of Beebee, Kail, and Wright that I consider in the current paper—all dated 2007 or later—are the only published attempts to reconcile the detail of Hume’s argument here with a New Humean interpretation (though there are some hints in Yolton, *Realism and Appearances*, 129–31).
  27. For previous presentations of this claim, see Peter Millican, “The Context, Aims, and Structure of Hume’s First *Enquiry*”, in *Reading Hume*, 27–65, at 58–60; “Against the ‘New Hume’”, 244–5, 252, n. 74; “Hume, Causal Realism, and Causal Science”, §8.
  28. Helen Beebee, “The Two Definitions and the Doctrine of Necessity”, *Proceedings of the Aristotelian Society* 107 (2007): 413–31. References in this section are to this paper, unless otherwise stated.
  29. I would also take issue with the appropriation of the term ‘real causal powers’ to signify *thick* causal powers, though Beebee’s usage here fits the assumptions of the position she is describing. According to the Old Humean



position, *real* causal power and necessity is, of course, to be understood in accordance with Hume's definitions.

30. As Beebee acknowledges (*Hume on Causation*, 176), the term 'natural belief' was coined by Normal Kemp Smith (*The Philosophy of David Hume*, 449, 454–8, 487–94) and is never used by Hume himself. Nevertheless, it has often featured in the literature of the New Hume debate, and has even been elevated into a 'doctrine' by some (e.g., Strawson, *The Secret Connexion*, 1–2, 13; 2000: 34; Buckle, *Hume's Enlightenment Tract*, 112, 211–12). Kemp Smith uses the term to cover 'two naturally conditioned (i.e., necessitated) propensities of the imagination' (p. 490), which respectively give rise to 'belief in continuing and therefore independent existence [of external objects], and . . . belief in causal dependence' (p. 455). But Hume gives quite different accounts of the mechanism behind these two propensities, and hence there is no unified 'theory of natural belief' *as an explanatory account*, beyond the suggestion that such belief is due to the operations of the imagination (as Beebee herself recognises at *Hume on Causation*, 201). In other Hume scholarship, the supposed theory of 'natural belief' has tended to play a rather different role, providing a focus for discussion of Hume's attitude of apparent *acceptance* or *endorsement* of certain naturally occurring beliefs, most contentiously the belief in God, irrespective of the mechanism by which they are generated (see Peter Millican, "Critical Survey of the Literature on Hume and the First *Enquiry*", in *Reading Hume*, 413–74, at 456–7, for a bibliographical overview). The best textual warrant for thus pairing together two 'natural beliefs'—in the external world and induction from experience (rather than 'causal dependence')—as potentially set in opposition to the belief in God comes from a 1751 letter of Hume to Gilbert Elliot: 'The Propensity of the Mind towards [the Design Argument], unless that Propensity were as strong & universal as that to believe in our Senses & Experience, will still, I am afraid, be esteem'd a suspicious Foundation. . . . We must endeavour to prove that this Propensity is somewhat different from our Inclination to find our own Figures in the Clouds, our Face in the Moon, our Passions & Sentiments even in inanimate Matter. Such an Inclination may, & ought to be controul'd, & can never be a legitimate Ground of Assent' (*HL* i 155).
31. Here I am very grateful to Helen Beebee for an extensive e-mail discussion which clarified details of the position she had in mind, which she takes to be the most plausible development of a sceptical realist approach (rather than an interpretative position to which she is personally committed). For more on this, see Beebee, *Hume on Causation*, 176–8, 201–4.
32. Indeed, Beebee's own reading of Hume's argument moves more in this direction than her gloss suggests.
33. Much of Beebee's exposition in her 2007 paper is structured around an itemisation of four readings of Hume's two definitions (pp. 417–19), of which only the first is Old Humean. However, in explaining how Hume's use of them can be reconciled with sceptical realism, she focuses on the second of the four (pp. 424–8), with only a brief mention of the last two (pp. 428–9). Here I follow her lead in ignoring nuances of difference between the New Humean approaches.
34. This is likely to be true for *any* detailed interpretation of the two definitions (including my own—Millican, "Hume, Causal Realism, and Causal Science", §4), and I intend no criticism of Beebee for speculatively developing Hume's very incomplete sketch of the relationship between the definitions. However, I find the emphasis she puts on the distinction between philosophical and natural relations (*Hume on Causation*, §4.6; "The Two Definitions and the Doctrine of Necessity", 418–19) unconvincing. Hume himself



introduces this distinction to clarify a now familiar ambiguity in the term ‘relation’: philosophers can consider things as ‘related’ by any number of arbitrarily invented relations, but this doesn’t mean that they’re related in the everyday sense (*T* 1.1.5.1). He immediately goes on to categorise the possible types of (philosophical) relation, later utilising this analysis in an attempt to identify those relations that are susceptible of a priori connection (*T* 1.3.1.1). After this, his only two (very cursory) mentions of the philosophical-natural distinction—in any of his works—are in single sentences at *T* 1.3.6.16 and *T* 1.3.14.31. Both are rather unclear, but seem to be saying little more than that causation can be thought of either as an abstract relation or as one that has particular relevance to human cognition in stimulating association of ideas. Thus it seems implausible to take the distinction as central to the interpretation of his two definitions, which are referred to repeatedly later in the *Treatise* (e.g., *T* 1.4.5.30–3, 2.3.1.4, 2.3.2.4), in the *Abstract* (*A* 26, 32), and in the *Enquiry* (*E* 7.29, 8.5, 8.25, 8.27).

35. See James A. Harris, *Of Liberty and Necessity: The Free Will Debate in Eighteenth-Century British Philosophy* (Oxford: Clarendon Press, 2005), chapter 2, and Peter Millican, “Hume’s Determinism”, in *Canadian Journal of Philosophy*, 40 (2010): 611–42, or for a brief summary, Millican, “Hume, Causal Realism, and Causal Science”, 704–6.
36. Cf. Millican, “Against the ‘New Hume’”, §1.2.
37. *Hume on Causation*, 177–8. This suggestion is standardly developed within the New Hume literature by appeal to the notion of a *relative* idea, as in Beebe, *Hume on Causation*, 177–9. But such a development, even if potentially applicable more generally (for doubts, see Millican, “Hume, Causal Realism, and Causal Science”, 658–9), would be of little help in this context, because Beebe’s reconciliation of Causal realism with Hume’s discussion of liberty and necessity crucially requires that it is the very idea of necessary connexion that Hume himself identifies (and not some surrogate ‘relative idea’) which itself represents the supposed objective reality.
38. Again I would emphasise (cf. note 31) that Beebe is not personally committed to this account, but proposes it as the best development of a New Humean position, the aim of her 2007 paper being to argue that *if* this New Humean approach is taken, *then* Hume’s argument concerning liberty and necessity poses no *new* difficulties for it.
39. For ease of exposition, I here ignore the caveat that only simple ideas need be directly copied from corresponding impressions; complex ideas can of course be put together from simples, rather than being directly copied.
40. In the *Treatise*, Hume anticipates a natural bias against his theory, arising from what seems to be a similar error: “Tis a common observation, that the mind has a great propensity to spread itself on external objects, and to conjoin with them any internal impressions which they occasion, and which always make their appearance at the same time that these objects discover themselves to the senses” (*T* 1.3.14.25). He likens this (*via* a footnote reference) to the error discussed at *T* 1.4.5.11–14, of supposing that the smell or taste of a fig is spatially coextensive with the fig itself. This error theory of mental spreading contrasts strongly with Hume’s far more positive view of moral taste, which ‘has a productive faculty, and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises, in a manner, a new creation’ (*Moral Enquiry* App 1.21). The latter ‘new creation’ appears to provide, for Hume, a *legitimate* standard of morality, which is not dismissed as erroneous, despite its apparently involving a similar spreading of internal sentiments onto natural objects. This contrast can be explained, I suggest, in terms of the *systematisation* involved in the ‘new creation’ of moral standards. To generate



such a legitimate standard, the imagination must be disciplined and constructive rather than just reactive. So a more appropriate comparison in the case of causal reasoning is with the formation of the ‘system of realities’ of the judgement, as described at *T* 1.3.9.3–4, and presumably involving the ‘rules by which to judge of causes and effects’ of *T* 1.3.15.

41. For more on this passage, which is treated as significant by Kail and Wright, see §V.
42. This claim (which Beebee adopts from Strawson, *The Secret Connexion*, 156, 159, 161–3) seems, indeed, to be entirely *ad hoc*, and without any textual basis, here or elsewhere. To illustrate just one of the difficulties, at *T* 1.4.2.36 and 42 Hume seems to countenance our ascribing the fictions of a perfect identity and a continued existence to interrupted images (cf. also *T* 1.4.2.52); likewise at *T* 1.4.3.3 and 1.4.6.6–7 he talks of our ascribing the fiction of perfect identity to a gradually changing object or a succession of related objects. If even a ‘fiction’—far less than a well-formed idea—is capable of being ‘ascribed’, then it is very hard to see why a supposed legitimate conception of ‘genuine’ thick necessity, if Hume thought we had one, would somehow resist it.
43. Though, as explained earlier, there would still be a question regarding the *uniqueness* of the understanding of necessity thus represented, and whether the same subjective idea of necessity could equally represent both *moral* and *physical* necessity, whilst still allowing a coherent distinction between them.
44. The most conspicuous exception to Kail’s claim is Strawson, *The Secret Connexion*, 13–15, 222–8, as cited in §I, but Wright, *Hume’s ‘A Treatise of Human Nature’*, 94, likewise talks of Hume’s ‘belief in the objectivity of power’ as ‘firmly implanted in human nature’.
45. “How to Understand Hume’s Realism”, 254. References in this section are to this work unless otherwise stated.
46. *Projection and Realism in Hume’s Philosophy*, 96.
47. See Millican, “Hume, Causal Realism, and Causal Science”, §6.
48. See Millican, “Hume, Causal Realism, and Causal Science”, §5.
49. See, for example, Wright, *Sceptical Realism*, 152–3; “Hume’s Academic Scepticism: A Reappraisal of His Philosophy of Human Understanding”, *Canadian Journal of Philosophy*, 16 (1986): 407–36, at 426–8; “Hume’s Causal Realism: Recovering a Traditional Interpretation”, in Read and Richman, 88–99, at 94–5; *Hume’s ‘A Treatise of Human Nature’*, 124–6; Broughton, “Hume’s Ideas about Necessary Connection”, 243–4, n. 27; “Our Aim in All Our Studies”, in Read and Richman, 198–210, at 201–3; and Kail, “Projection and Necessity in Hume”, 45–46; *Projection and Realism in Hume’s Philosophy*, 107–8, all of whom prominently cite *T* 1.4.3.9.
50. For more on this passage, see the discussion of Wright’s view in §VI.
51. John Locke, *An Essay Concerning Human Understanding*, edited by P. H. Nidditch (Oxford: Clarendon Press, 1975, originally published 1690).
52. Kail’s statement (p. 265) that ‘we are already told at the beginning of [Section 8] that disputes at the metaphysical level can never arrive at a determinate conclusion’ thus involves a clear misreading. Hume tells us only that disputes *which lie entirely beyond the reach of human capacity* can never arrive at a determinate conclusion.
53. For detailed substantiation of this claim, see Millican, “Hume, Causal Realism, and Causal Science”, §8 (which extensively quotes parallel passages from the *Treatise* and *Enquiry*) and “Humes Old and New”, §VIII (which cites parallel passages from all three versions).
54. After this quotation, but before going on to discuss the violation issue, Kail makes two additional points, first denying that Hume’s definitions of cause can be intended as exhaustive analyses of the meaning of ‘cause’ because



they are not equivalent, and then quoting *E* 8.21 to highlight Hume's use of epistemic language 'in the midst of the discussion'. Interpretation of the two definitions is notoriously tricky, and too big an issue to be addressed here, but see Millican, "Hume, Causal Realism, and Causal Science", §4, for my own view. As for *E* 8.21, by this stage Hume has completed his main argument—we are no longer 'in the midst of the discussion'—and he is discussing an error theory of why others *misunderstand* necessity, based largely on epistemic considerations. Moreover, he is using this paragraph to set up a refutation of the speculation that he imputes to his opponents, a refutation which is solidly conceptual rather than epistemological, based on the semantic limits of our ideas (a brief sketch of all this was given in §III).

55. This correspondence is very evident in the texts. Indeed, *Enquiry* 8.19 is virtually identical to the penultimate paragraph of *Treatise* 2.3.1, while *Enquiry* 8.21 and *Treatise* 2.3.2.1 both embark on the question of why people have been so inclined to deny the doctrine of necessity.
56. Here I draw on a very helpful e-mail discussion with John Wright, for which I am grateful.
57. Along with *T* 2.3.1.16, which is mentioned in one of the following quotations, these are *T* 1.3.9.10, which says that in the communication of motion by impulse, the resemblance of cause and effect can 'make us imagine them to be absolutely inseparable', and *T* 1.3.11.4, which says that we form such a 'habit of surveying [constantly conjoined objects] in that relation [of cause and effect], that we cannot without a sensible violence survey them in any other'.
58. Again, see Millican, "Against the 'New Hume'" and "Hume, Causal Realism, and Causal Science", for discussions of Hume's alleged Causal realism that are relatively comprehensive, though they do not much address nuances of variation between the different New Humean accounts.
59. Wright also draws on work of my own (cf. "Hume's Determinism") to make the point that Hume seems indeed to be identifying his own view with what he has described as 'universally acknowledged' (pp. 177–8).
60. See §IV, just before note 32.
61. This commitment is very clear in the *Treatise*, the *Abstract*, and the *Enquiry*, with the Copy Principle described as the 'first principle' of the *Treatise* (*T* 1.1.1.12), the 'first proposition' of the *Abstract* (*A* 6), and trumpeted as a major discovery in all three works (*T* 1.2.3.1, *A* 7, *E* 2.9). Hume's arguments for the principle are uncharacteristically weak (cf. Jonathan Bennett, "Empiricism about Meanings", in *Reading Hume*, 97–106, at 99–103), but there is no evidence of his having recognised this, perhaps because—as he acknowledged in a 1763 letter to Thomas Reid—his empiricist principles of the 'way of ideas' were so firmly entrenched as 'the common ones' (*HL* i 376), presumably due to Locke's influence. It is notable that in a letter the previous year to Hugh Blair, responding to Reid's draft *Inquiry*, Hume specifically defended his arguments for the Copy Principle, which Reid had apparently alleged to be 'not supported by any Colour of Argument' (Thomas Reid, *An Inquiry into the Human Mind on the Principles of Common Sense*, edited by Derek R. Brookes [Edinburgh: Edinburgh University Press, 1997, originally published 1764], 257).
62. §1 of Millican, "Hume's Determinism", argues a similar point against James A. Harris, "Hume's Reconciling Project and 'The Common Distinction betwixt Moral and Physical Necessity'", *British Journal for the History of Philosophy*, 11 (2003): 451–71, and *Of Liberty and Necessity*, who suggests an inconsistency between thoroughgoing determinism and Hume's views on induction and causation.



63. This vital point—which applies equally to Hume’s argument that ‘matter and motion’ may be ‘the causes of thought’ (*T* 1.4.5.30–33)—is emphasised in §§7–10 of Millican, “Hume, Causal Realism, and Causal Science”. Hume’s analysis of causation is standardly interpreted as *sceptical*, but in fact his application of it is scientifically constructive. By making causal necessity a matter of constant conjunction (rather than something mysterious), Hume is able to argue for its thoroughgoing application to the mental world, in a way that he could not if it were metaphysically ‘thick’.
64. Again, I am not suggesting that there are no *philosophical* difficulties in the Old Hume position, for example, in making sense of the ascription of necessity to objects based on a subjective impression (cf. Millican, “Against the ‘New Hume’”, §3.5). The point is that these are *Hume’s* own philosophical difficulties: problems that he himself clearly encounters.
65. Note that Hume implicitly relies on a distinction between *absolute* (or conceptual) possibility, which is a priori and revealed by his Conceivability Principle, and *causal* possibility, which cannot be discovered through mere conceivability and depends on what the causal laws happen to be. Often Hume’s language is ambiguous between the two (and sometimes also *epistemological* possibility), but in the crucial discussions he mostly uses ‘possible’ in the absolute sense (especially when applying the Conceivability Principle), and ‘necessary’ in the causal sense (especially in his treatment of necessary connexion and determinism). See Millican, “Hume, Causal Realism, and Causal Science”, 676–7.
66. For more on all this, see §§9–10 of Millican, “Hume, Causal Realism, and Causal Science”.
67. For example, Peter Kail’s work on projection and realism—though it seems largely to have taken off from his interest in the New Hume debate—has now gone well beyond it, exploring connections between quite different aspects of Hume’s philosophy, including the external world, personal identity, morality, and religious belief.
68. My recent research on this chapter has benefited from the hospitality of the Institute for Advanced Studies in the Humanities (IASH) at the University of Edinburgh, in the role of Illumni David Hume Fellow. I am very grateful both to the Edinburgh Illumni and to IASH for providing such a delightful and stimulating context for my work.



## 8 Pouring New Wine into Old Skin

### The Meaning of Hume's Necessary Connexions

*Constantine Sandis*

Neither do men pour new wine into old wineskins. If they do, the skins will burst, the wine will run out and the wineskins will be ruined. No they pour new wine into new wineskins, and both are preserved.

Matthew 9:17.

#### I. ANOTHER SIDE OF DAVID HUME

The bulk of secondary literature on Hume up until the turn of the century depicts him as holding the *regularity* theory of causation according to which causal relations in the 'external world' amount to no more than regular and/or constant conjunction.<sup>1</sup> It is no surprise, then, that so-called New Humeans, who argue that Hume couples sceptical epistemology with realist metaphysics, choose to illustrate the 'standard view' they wish to attack through texts that are paradigmatic of this approach:<sup>2</sup>

If Hume is right . . . even if God were to look at [two causally related] events, he would discern nothing relating them other than that one succeeds the other.<sup>3</sup>

Hume's conclusion [is] that so far as the external objects which are causes and effects are concerned there is only constant conjunction . . . regularity and constant conjunction are all that exist.<sup>4</sup>

All that ever happens in the world independently of minds is that one thing succeeds another and resembles other instances that followed similar antecedents.<sup>5</sup>

If New Humeans are to be believed, this regularity account of causation (often misleadingly labelled 'anti-realist' instead of 'reductivist') is to be contrasted with an account of causation that involves necessity. And yet the term 'necessity' is conspicuous by its absence in passages such as those quoted earlier. For all that has been quoted, Hume might be working with a notion of causal necessity that is itself reducible to constant conjunctions



between natural regularities and mental impressions. In fact, many of the ‘old’ come close to saying just this, albeit with a tendency to overemphasise the role of the mind as the *residence* of necessary connexions, thus making an otherwise cogent reading vulnerable to attack:

... causal necessity is really in the mind of the observer and not, as the uninstructed laity would have it, in whatever “objects” are said to be causally related.<sup>6</sup>

Hume’s ‘main conclusion’ on the topic is that ‘necessity is in the mind, not in objects’.<sup>7</sup>

... both on Hume’s view and on the view taken here, causal necessity is not to be found “*in nature*”. In nature there are only regular sequences.<sup>8</sup>

Necessities do not exist here in the world; but only in our ways of describing it.<sup>9</sup>

On such readings Hume is an error theorist about necessity without being an antirealist about it. This is because the error he identifies is not that of believing in the *existence* of necessity (or our knowledge of it) but, rather, an error of *location*: we project onto the external world what in actuality resides only in our minds.

Necessity so conceived will count as ‘real’ necessity for the metaphysically inclined, but interpretations of Hume should not be mischaracterised—let alone disqualified—on such grounds. Having mischaracterised error theorist interpretations of Hume as ‘anti-realist’, New Humeans suggest we would do better to instead think of him as a *sceptical realist* best read as claiming that (i) there *are* necessary connexions in nature even though (ii) we could never know for certain that they exist, since we can neither perceive them nor offer an *a priori* proof. It is argued that older interpretations made the mistake of conflating Hume’s epistemology (said to be sceptical) with his metaphysics (said to be realist). After all, it would be inconsistent for Hume the sceptic to pronounce that we *know for certain* that there are no necessary connexions in nature, besides which he generally maintains that human nature prevents us from abandoning beliefs that reason cannot justify (cf. T 130ff.).

New Humeans offer a persuasive argument against a regularity interpretation according to which Hume is making a negative existential statement concerning necessity in the ‘external’ world. Such a thesis should, however, be distinguished from a regularity view which incorporates an account of necessity (projectivist or otherwise); *some* of the older interpretations of Hume clearly nod towards such a view, albeit whilst contrasting between regularity and necessity when no such contrast is to be found in Hume’s



texts. While by no means blind to Hume's empiricist account of meaning (according to which the meaning of any idea is fixed by corresponding impressions), both Old and New Humeans fail to take it seriously, possibly under the influence of the wishful thought that Hume is interested in the nature of causality and/or causal knowledge, over and above that of causal reasoning. Hume nowhere claims that causation consists of mere regularities. To anticipate, he takes the notion of necessity to be 'conformable to the two definitions of cause, of which it makes an essential part' (T 409; cf. EHU 60).<sup>10</sup>

## II. SELF-PORTRAITS

Kenneth Winkler has sketched a similar picture of Hume:

Hume need not say that there is no such thing as objective connection; it is enough for him to say that we cannot *in any way* conceive of it, and that as a result we cannot believe in it. According to Hume, our only causal conceptions (of any sort) are captured either by his own definitions, or by the loose and inaccurate idea against which he argues. And we can avoid the latter conception, even when we step outside the study.<sup>11</sup>

Immediately after, however, he adds that on Hume's view '[n]ecessity *as we understand it* lies entirely in the mind. We cannot even *think* or *wonder* about it as it exists in objects'.<sup>12</sup> This further comment misleads in two different ways. First, the qualifier 'as we understand it' follows New Hume in suggesting that Hume can allow for talk of necessity in any other sense than that which we may understand. Secondly, the projectivist suggestion that for le bon David (this) necessity 'lies entirely in the mind' runs counter to the fact that Hume describes necessity as 'an intelligible quality' which he ascribes 'to matter' (T 410), is also present 'in human actions' (T 411), and repeatedly claims that causes and effects are *objects* (e.g., T 93 & EHU 59 & 127). This hardly resembles a quality that lies in the mind (entirely or otherwise).

To say that necessity exists *in any given object* is to either make a claim about constant conjunction or the determination of the mind; there is no question of it not being attributable to matter here. Moreover, as Harold Noonan<sup>13</sup> and Henry Allison<sup>14</sup> both point out, Hume generally takes (but cf. T 156 & 165) 'determinations' of the mind to be propensities or dispositions to behave in a certain ways under certain conditions, making them implausible candidates for impressions (which he classifies as perceptions). Hume *does* talk of an impression of necessity in the mind but given his belief that (a) all ideas are derived from impressions and (b) all impressions are in the mind, this hardly marks necessity as belonging to the mental realm in any unique sense.



Allison persuasively argues that it is crucial to distinguish between (a) the question of whether Hume's *concept* of causation includes a notion of necessity and (b) the question of whether Hume is a (metaphysical) realist about causation.<sup>15</sup> He answers (a) positively (Hume's 'concept of a cause involves a necessary connection with its effect') and (b) negatively (Hume believes that 'we have no grounds for positing anything like necessary connections in nature').<sup>16</sup> I part company with Allison in his response to (b), taking Hume's preferred concept of necessity to involve nothing that we do not experience on a daily basis. By contrast, Allison sees Hume as being in complete agreement with Malebranche not only in insisting that 'genuine causation requires a necessary connection with its effect' (which he takes Hume to understand as logical necessity<sup>17</sup>) but also in arguing that 'causation, so construed, was not to be found in nature'.<sup>18</sup>

I reject the latter part of Allison's reading, but not the former. Kenneth Clatterbaugh, by contrast, dismisses both, claiming that while 'Hume agrees with Malebranche that necessary connexion is a crucial part of our idea of cause' (cf. T 77), he nonetheless maintains that 'Malebranche's mistake lies in his insistence that a true cause is necessarily connected to its effect'. Clatterbaugh claims this is a consistent position because on Hume's view 'the necessary connexion between cause and effect is neither logical necessity nor metaphysical necessity but a psychological necessity derived from custom and habit',<sup>19</sup> as seemingly supported by the following passage:

Upon the whole, necessity is something that exists in the mind, not in objects; nor is it possible for us to ever form the most distant idea of it, consider'd as a quality of bodies. (T 165–6; cf. EHU 59)

The remark is unlikely to convince realist interpreters (sceptical or otherwise) who would rightly point out that (i) the qualification 'upon the whole' points to exceptions of one kind or another and (ii) Hume does not here deny that necessity is a bodily quality but only that we cannot form an idea of it as such. Even those who agree with the spirit of Clatterbaugh's interpretation may still reject its letter, for Hume's treatment of the problem of induction arguably reveals that he followed his deist predecessors in failing to properly distinguish between these and other kinds of necessity.<sup>20</sup> Moreover, Clatterbaugh's reading is also difficult to square with Hume's repeated claim that causes and effects are *objects*.

Paul Russell argues that it is a mistake to try to iron out such tensions, suggesting that the schizophrenic thread which runs throughout Hume's writings is no accident. According to Russell, Hume's sceptical realism regarding the external world leads him to embrace a double ontology which directly motivates his two definitions of a cause (physical and mental).<sup>21</sup> Hume, so interpreted, believes that for all we know we only ever refer to mental impressions (but cannot help thinking that the world we experience lies outside our minds). Our prephilosophical belief in the *continuous* and



*distinct* existence of objects is accordingly said to be the result of neither reason nor the senses but of a *natural instinct* which combines with the imagination to resolve cases of apparent conflict (T 1.4). Such beliefs have as little to do with a metaphysical outlook as, say, a dog's belief that its food is in the bag (head-burying ostriches do not espouse relativist theories of perception).

Causation, on such an outlook, only lies in the mind in the sense in which billiard balls reside there too (we would do well to ask ourselves what meaning it allows Hume to give to the term 'external world'). Hume even appeals to his empirically defined notion of a necessary connexion to distinguish between *accidental* regularities and *causal* ones. Whether or not he can do so successfully while defining physical necessity as 'the constant conjunction of objects, along with the determination of the mind (T 171)', is another matter.

Hume's account of necessary connexions is both simpler and more coherent than is commonly imagined. That it should strike us as implausible is hardly surprising given that philosophers have (a) always had a taste for dichotomies between realist and antirealist positions and (b) long abandoned the Copy Principle. Too many scholars are quite literally projectivists, superimposing their own views onto Hume's text in the name of charity.<sup>22</sup> But it is misguided to use philosophical plausibility as a guide to charity, not least because philosophical plausibility is so rarely uncontentious. The end result will almost inevitably be a self-portrait of the interpreter projected onto rough sketches of the interpretee.

### III. ENQUIRY 61 REVISITED

Of equal importance is the fact that the interest in causation in contemporary metaphysics and philosophy of science is very far removed from Hume's own concerns. In both the *Treatise* and the two *Enquiries* (and much of the *History of England*), Hume's project is that of understanding human nature—how and why it is that we humans feel, think, reason, and act as we do: what *motivates* our thought and action? His discussion of causation must thus be approached in terms of this wider investigation.

One common tactic is to assume that Hume is primarily investigating why, how, and when it is that we acquire our concept of causation and that anything he says (if he says anything at all) about what causation *is* (what it does or does not consist in) has the purely instrumental purpose of shedding light on how the very nature of causality may somehow give rise to our beliefs about it. While not entirely inappropriate, such an approach is misleading in two different ways. First, it implies that Hume could allow that we might have causal knowledge that transcends any ideas derived from impressions. Secondly, it suggests that his chief interest in causation results from the persuasion that facts about our causal ideas are *directly* relevant



to questions relating to human nature. Yet Hume tells his readers quite explicitly that it is the *idea* or *concept* of causation that he is investigating, and that his interest in this arises indirectly out of a more general interest in the way(s) in which human beings naturally reason:

'Tis evident, that all reasonings concerning *matters of fact* are founded on the relation of cause and effect . . . In order therefore to understand these reasonings, we must be perfectly acquainted with the idea of a cause; and in order to do that, must look about us to find something that is the cause of another. (T 649)

We must reflect on our *idea* of a cause in order to better understand our reasonings. Hume tells us that our idea of causation is a *complex* one, made up of (at least) the following three simple ideas: (i) contiguity; (ii) priority; (iii) necessary connexion (T 75–77). He quickly locates *impressions* for the first two, but has trouble with the third. It follows from the Copy Principle (T 1–7 & EHU 61) that in cases where we believe we have an idea of *x* but cannot find a corresponding impression, the terms we use to try to give expression to our thoughts have no *clear* meaning:

. . . when we speak of a necessary connection between objects and suppose that this connection depends upon an efficacy, or energy, with which any of these objects are endowed; in all these expressions, so apply'd, we have really no distinct meaning, and make use only of common words, without any clear and determinate ideas. (T 162)

Hume immediately adds that this is not because the expressions 'never have any meaning' but that they lose their meaning by being 'wrong apply'd', a meaning he will later rediscover by locating the relevant impression, hence the italicised 'seems' in the parallel passage of the first *Enquiry*:

We have sought in vain for an idea of power or necessary connexion in all the sources from which we could suppose it to be derived . . . the necessary conclusion *seems* to be that we have no idea of connexion or power at all, and that these words are absolutely without any meaning, when employed either in philosophical reasonings or common life. (EHU 58)

We may say . . . that the vibration of this string is the cause of this particular sound. But what do we mean by this affirmation? We either mean *that this vibration is followed by this sound, and that all similar vibrations have been followed by similar sounds*: Or, *that this vibration is followed by this sound, and that upon the appearance of one the mind anticipates the senses, and forms immediately an idea of the other*. We may consider the relation of cause and effect in either of these two lights; but beyond these, we have no idea of it. (EHU 60)



A. J. Ayer has objected that Hume's empiricism actually weakens his case for this view:

Hume has failed to do justice to his case by making it appear to rest on an empirical generalization, rather than a logical argument . . . he is right in saying that no relations of force or power can be detected in the actual phenomena. The point which he overlooks is that it would not matter if they were detectable . . . for [this would] be entirely neutral with regards to what happens at any other time and place . . . on this interpretation, what looked like the empirical proposition that necessary connexion is not detectible in any single instance of a conjunction of matters of fact, is promoted into a logical truth. For the relation is now defined in such a way that we have to examine every single instance of the phenomenon in question, in order to discover that it holds in any one of them.<sup>23</sup>

Ayer's point is best cast in terms of Simon Blackburn's distinction between a *causal nexus* (that may guarantee necessity at any one particular place and time) and a *straitjacket* on this nexus, which ensures that it remains there in the future.<sup>24</sup> Even if, *per impossibile*, we could observe a nexus, no finite number of observations could ever possibly give rise to the impression of a straitjacket. That is to say, even if we could somehow observe something in *x* that causally necessitates that *y* will follow in this particular instance, this would not demonstrate that this or any other *x* would necessarily be followed by this or any other *y* in any otherwise identical future circumstances. It is perverse, however, to offer this distinction as a helping hand to Hume, for if we have no impression of a straitjacket we can have no Humean idea of it either.

It may be thought that Hume's strategy faces an insuperable difficulty here: is it not impossible for us to know what impression we are looking for if we cannot have any distinct idea of it until we have located the impression in question? But the problem is merely apparent, for we know what the idea is meant to be *about*, namely, an observed connection(s) between any two or more given objects. This alone will fix the content of whatever ideas it gives rise to.

#### IV. IT'S ALL OVER NOW, SHADE OF BLUE

It is instructive to compare the (temporarily) missing impression of a necessary connexion to the case of the missing shade of blue. In the *Treatise*, Hume simply shrugs his shoulders and states that this type<sup>25</sup> of instance is too slight an exception to warrant the revision or abandonment of the Copy Principle:



[T]his may serve as proof that the simple ideas are not always, in every instance, derived from the corresponding impressions; though this instance is so singular, that it is scarcely worth our observing, and does not merit that for it alone we should alter our general maxim. (EHU 21–2; cf. T 6)

There immediately follows a striking new paragraph in which Hume proceeds to talk of jargon that has ‘drawn disgrace’ upon ‘metaphysical reasonings’ and of how the Copy Principle can help us distinguish between proper and improper uses of philosophical terms (which are particularly suspect in this regard), the latter being employed ‘without any meaning or idea’:

Here, therefore, is a proposition, which not only seems, in itself, simple and intelligible; but, if a proper use were made of it, might render every dispute equally intelligible, and banish all that jargon, which has so long taken possession of metaphysical reasonings, and drawn disgrace upon them. . . . When we entertain, therefore, any suspicion that a philosophical term is employed without any meaning or idea (as is but too frequent), we need but enquire, from what impression is that supposed idea derived? And if it be impossible to assign any, this will serve to confirm our suspicion. By bringing ideas into so clear a light we may reasonably hope to remove all dispute, which may arise, concerning their nature and reality. (EHU 22)

The text includes a lengthy footnote in which Hume accuses Locke of misusing the words ‘innate’ and ‘idea’.<sup>26</sup> His chief point is that, despite the missing shade counterexample, the Copy Principle itself is not only intelligible but could itself be used as a criterion for separating intelligible propositions from unintelligible ones. Given that the principle is offered as a matter of fact (rather than a principle concerning the relation of ideas), Hume must allow that it is *logically possible* that ideas are not derived from impressions.<sup>27</sup> As Peter Kail puts it, ‘The Copy Principle is an a posteriori and contingent generalization’.<sup>28</sup> Kail explains that ‘[a]s such, it is unable to deal with possible counter-examples’<sup>29</sup> despite the fact that Hume admits the counterexample of the missing shade of blue. It is important to ask, however, why Hume thinks that his generalisation should *not* be abandoned at the drop of a missing shade and, more importantly, why he refuses to adopt a similarly nonchalant attitude towards the idea of necessary connection.

The answer is that while Hume considers the notion of the shade to be a simple one, he suspects that the terms ‘cause’ and ‘necessary connexion’ are all too frequently used in a loose and unhelpful senses. The worry is not that they could *never* have a legitimate use, but that neither philosophers nor the ‘vulgar’ seem to adhere to one, Hume suspecting that philosophical controversy arises from linguistic vagueness and ambiguity when ‘a few intelligible definitions would immediately have put an end to the whole



controversy' (EHU 63). In the case of causation and necessity, not only have philosophers failed to affix the *same* ideas to the same terms but, given the serious lack of relevant impressions, they have failed to affix *any* ideas to causal terms which are thereby employed with no clear meaning (cf. T 162 & EHU 59, quoted in §III).

While Kail may be right to claim that the Copy Principle 'is a genetic claim, a claim about the source of ideas, which is couched in terms of psychological fact' and that 'showing the origin of a thought . . . seems independent of its meaning',<sup>30</sup> it is far from clear that *Hume* thinks that origin and meaning are independent. Kail is also right to point out that Hume has no *theory* of meaning, certainly not in the sense in which Grice, Quine, Davidson, or even Locke do.<sup>31</sup> Be that as it may, he has a *view* of meaning and it is empiricist to the extreme.

## V. POSITIVELY HUME'S MEANING

Some find it anachronistic to parse Hume's passages on meaning as we would any similar-looking remarks in post-Fregean philosophy. Galen Strawson, for example, writes as follows:

Hume—with Locke, Berkeley, and many others—uses the word 'unintelligible' in the literal sense which survives in the standard non-philosophical use of the word—as when we say that a message is unintelligible, meaning simply that we cannot understand it although it exists . . . When Hume says that something is unintelligible, then, he means that we cannot understand it. In particular he means that we cannot form an idea of it or term for it that has any positive descriptive content on the terms of the theory of ideas. To say this, however, is not to say that we cannot refer to it, or that the notion is incoherent . . . Our understanding of terms like 'meaning' and 'unintelligible' is not the same as Hume's. There is obviously no difficulty in the idea that we may successfully use a term to refer to something which has some manifestation in our experience, even though we have no positive conception of its nature, over and above the thought that it is something and has the manifestation that it has.<sup>32</sup>

Strawson's claim is that Hume can allow that we can form an idea of something whose nature we (lacking the relevant impressions) cannot understand. Its content would thus be constrained by whatever experiences give rise to the possibility of something whose nature remains secret to us, a move motivated by Hume's notorious references to nature's secret powers (see §VII).<sup>33</sup> Peter Kail constructs such a notion by appealing to the predicted effect of our becoming acquainted with such secret things. He dubs this the 'Bare Thought necessary to meet the semantic threat to realism', explicating further:



We can specify uniquely that which we cannot understand (causal power) by saying that it is that feature that, were we acquainted with it, would yield *a priori* inference and render it inconceivable that the cause would not be followed by its effect.<sup>34</sup>

But even if this is so, the linguistic evidence suggests that this is far from what Hume was getting at. The following passage from Cudworth, for instance, illustrates it would have been perfectly natural for Hume to think that unintelligibility (and inconceivability) entailed *impossibility*, thus casting a doubt at the suggestion that Hume could have referred to powers that are unintelligible to us without worry:

Because this Infinite Power, is a thing, which the Atheists quarrel much with, as if it were altogether Unintelligible and therefore Impossible, we shall here briefly declare the Sence of it, and render it (as we think) easily Intelligible or Conceivable, in these Two following steps. First, that by Infinite Power is meant nothing else, but Perfect Power, or else as Simplicius calls it, ‘oli dinamis’, a Whole and Entire Power, such as hath no Allay and Mixture of Impotency, nor any Defect of Power mingled with it . . . all whatsoever is Conceivable, and which does not imply a Contradiction; for Conception is the Only Measure of Power and its Extent; as shall be shewed more fully in due place.<sup>35</sup>

Linguistic evidence further confirms that we have no reason to think that Hume’s usage of the word ‘meaning’, at least as it relates to words and terms, is different from ours. The year before the *Enquiry* was published, Samuel Johnson published his *Plan of a Dictionary of the English Language*, the prospectus in which he publicly announced his intent to ‘ascertain the meaning of our English idiom’ by considering ‘the words and phrases used in the general intercourse of life, or found in the works of those whom we commonly style polite writers’. He states that in this way ‘words are illustrated in their different significations’ for the benefit of those, such as the unlearned who ‘much oftener consult their dictionaries for the meaning of words, than for their structures or formations’, adding that ‘the words that most want explanation are generally terms of art; which, therefore, experience has taught my predecessors to spread with a kind of pompous luxuriance over their productions’. When discussing spelling, Johnson notes that sometimes ‘by the change of one letter or more, the meaning of a word is obscured, as in *farrier* for *ferrier* . . . *gibberish* for *gebrish*’. He also talks of ‘the comparison of one meaning with another’ of ‘synonyms’, of words ‘signifying’ things and of the ‘signification in words’. Johnson gives examples of words in which ‘the strict and critical meaning ought to be distinguished from that which is loose and popular’, speaks of ‘accidental’ and ‘adventitious’ meanings and of it being ‘necessary to sort the several senses of each word, and to exhibit first its natural and primitive signification’.



It is hard to see how Johnson could have meant anything else by the meaning or signification of words than what we currently mean. Officially in the business of explaining the meaning of words, he feels no need to ever give an explanation of what he means by ‘meaning’ (though we know the answer from the entries just listed). While he was arguably the first to compile a *purely* descriptive dictionary, namely, one whose *sole* purpose is that of making explicit what would have already been known by someone who understood the word correctly (as opposed to helping the reader extend or correct his or her vocabulary), Bailey’s *Universal Etymological Dictionary* (1721) contained some such attempts.<sup>36</sup> So, while it may be anachronistic to read Berkeley’s claims about the meaning of ordinary words (in 1710’s “Treatise Concerning the Principles of Human Knowledge”) as explicit statements of his implicit linguistic knowledge,<sup>37</sup> by the time of Hume’s *Treatise* (1739–40), let alone his first *Enquiry* (1748), the idea of giving the *meaning* of words in this sense could not have been so foreign.

It is true that Johnson’s dictionary defines ‘intelligibility’ as ‘possibility to be understood’, intelligible as ‘to be conceived by the understanding; possible to be understood’, ‘unintelligible’ as ‘not such as can be understood’, and ‘unintelligibly’ (with quotations from Locke) as ‘in a manner not to be understood’. But Johnson also quotes the following from Dryden:

Something must be lost in all translations, but the sense will remain, which would otherwise be maimed, when it is scarce *intelligible*.<sup>38</sup>

This suggests that an unintelligible translation is one where the *sense* is lost. Johnson lists three synonyms for ‘sense’: understanding, meaning, and import. The first allows that a senseless term or expression may refer in a way in which the other two don’t seem to. Johnson might allow for a sense of ‘meaning’ that equates it with understanding, but he also relates it to *signification*:

The act of making known by signs (South.); meaning expressed by a sign or word (*Holder*); To declare by some token or sign (Shakespeare).<sup>39</sup>

The notion resembles our current concept of meaning when applied to words or sentences. Moreover, Hume does not only talk of *ideas* being unintelligible, he also states that *terms* and *words* are meaningless when used in a certain way. Could he really have meant that we do not *understand* what their sense is while believing that they have an (unintelligible) reference, that a word that fails to signify can still succeed in referring? Surely this suggestion is more anachronistic than the alternative.

Hume also presents us with strongly felt *philosophical* reasons for denying that the Bare Thought has any meaningful content. First, there is the question of whether Hume’s assertions concerning *secret* powers are



capable of being true or false if we understand them as Strawson and Kail suggest we ought to. Hume's Fork presents us with a notoriously narrow account of what he is willing to allow here (T 458; cf. 413–4). The Bare Thought takes the form of a possible fact yet cannot qualify as such on Hume's understanding of such things, for it makes no claims about a *real*, empirically verifiable, existence. Consider Hume's infamous rejection of schools of divinity and metaphysics:

If we take in our hand any volume; of divinity or school metaphysics, for instance; let us ask, Does it contain any abstract reasoning concerning quality or number? No. Does it contain any experimental reasoning concerning matter of fact and existence? No. Commit it then to the flames: for it can contain nothing but sophistry and illusion. (EHU 165)

Upon asking Hume's questions in relation to the Bare Thought, it becomes clear that his aforementioned lack of interest in metaphysical concerns is no mere preference of enquiry. Whatever *we* may think of the metaphysical disputes about causation, they have no place in Hume's system. The semantic threat may a threat to *sceptical* Realism and *quasi*-Realism, but not to the deflationary view I am here attributing to Hume.

## VI. TIGHT CONNEXION

Whether he takes the Bare Thought of a 'necessary connexion' to be meaningless does not ultimately matter much, for Hume announces that 'there still remains one method of avoiding this conclusion, and one source which we have not yet examined',<sup>40</sup> and proceeds to locate that elusive impression and define 'necessity' in accordance with it. The definition he gives is far removed from that given earlier by Kail, acquaintance with what he re-baptises 'causal necessity' being compatible with our ability to conceive that the same cause might have not been followed by its effect, even if we can't help but to *expect* that it will (hence the *psychological* problem of induction).

Recalling his account of how it is that we come to make causal claims, Hume explicitly states that our idea of causation *includes* that of necessity (which is what gave rise to the puzzle in the first place):

When a particular species of event has always, in all instances, been conjoined with another, we make no longer any scruple of foretelling one upon the appearance of the other, and of employing that reasoning, which alone can assure of any matter of fact or existence. We then call the one object, *Cause*; the other *Effect*. We suppose that there is some connexion between them; some power in the one, by which it infallibly



produces the other, and operates with the greatest certainty and strongest necessity. (EHU 60)

He next proceeds to (i) locate the missing impression in our minds and (ii) suggest that *all we mean* when we say that two objects are connected is that their constant conjunction gives rise to such thoughts:

It appears, then, that this idea of a necessary connexion among events arises from a number of similar instances which occur of the constant conjunction of these events . . . after a repetition of similar instances, the mind is carried by habit, upon the appearance of one event, to expect its usual attendant, and to believe that it will exist. This connexion, therefore, which we feel in the mind, this customary transition of the imagination from one object to its usual attendant, is the sentiment or impression from which we form the idea of power or necessary connexion. Nothing farther is in the case. This is the sole difference between one instance, from which we can never receive the idea of connexion, and a number of similar instances, from which it is suggested . . . When we say, therefore, that one object is connected with another we mean only that they have acquired a connexion in our thought, and give rise to this inference, by which they become proof of each other's existence. (EHU 59)

According to Mackie, Hume is here 'telling us what we can properly mean rather than what we ordinarily mean', for he 'thinks that the ordinary meaning is itself mistaken, and calls for reform'.<sup>41</sup> By contrast, I have been urging that Hume believes that what we can properly mean is *all* that we can ever mean (ordinarily or otherwise). Meanings are either present or absent, precise or vague; they cannot be correct or mistaken, save for the sense in which 'correct meaning' simply means conventional meaning. Hume's semantic revisionism is an attempt to capture a meaningful application of an otherwise empty expression, one whose roots lie in the fact that the constant conjunction of *x* and *y* causes us—in the *very sense* of 'cause' being defined—to form expectations that give rise to inductive inferences.

Hume thereby believes that 'the constant conjunction of objects determines their causation' (T 173).<sup>42</sup> Indeed, as Millican points out, he states 'with conspicuous frequency' that the nature of any impression of necessary connexion will determine 'the very essence' of necessity.<sup>43</sup> John P. Wright assumes that Hume generally uses 'determine' as a synonym for 'discover', thereby talking of Hume's 'central discovery concerning causation', concluding that 'while Hume scholars agree on some of the details there is general agreement that this discovery was epistemological'.<sup>44</sup> But Hume makes no discovery, epistemological or metaphysical, about the nature of causation. As noted in §IV, his aim is to elucidate the concept with the aim of putting



to rest age-old disputes caused by ‘some ambiguous expressions, which keep the antagonists still at a distance, and hinder them from grappling with each other’ (EHU 63) for the simple reason that they ‘affix different ideas to the terms employed in the controversy’ (EHU 62). In the case of ‘liberty and necessity’, Hume hopes to ‘put an end to the whole controversy’ through ‘a few intelligible definitions’ (EHU 63), most centrally those of ‘liberty’ and ‘necessity’. If there are discoveries here they are empirical discoveries relating to semantics, heuristics, and psychology.<sup>45</sup> Wright claims that ‘what Hume scholars mainly disagree about is his account of what causal power or causal connection really is—what we may call the *ontology* of causation’,<sup>46</sup> in particular whether or not it amounts to anything more than constant conjunction. By contrast I have tried to show that when Hume talks of the ‘the very essence’ of causation or necessity, his discoveries are limited to empirical observations relating to the impressions that underlie linguistic usage.

Hume does declare that we ‘feel’ a connexion in the mind but the connexion we feel is not the necessary connexion but the impression which gives rise to the idea of such a connection, besides which Hume in any case claims that *all* beliefs are more properly felt than judged of:

Belief is nothing but a peculiar feeling . . . or sentiment. . . . ’Tis felt rather than conceived, and approaches the impression, from which it is deriv’d, in its force and influence . . . it is something *felt* by the mind which distinguishes the ideas of judgments from the ideas of the imagination. (T 624)

Hume’s own resolution thereby appeals to a clear and distinct definition of necessity, as stated at the end of the *Enquiry* section I quoted from previously:

Our idea, therefore, of necessity and causation arises entirely from the uniformity observable in the operations of nature, where similar objects are constantly conjoined together, and the mind is determined by custom to infer the one from the appearance of the other. These two circumstances form the whole of necessity, which we ascribe to matter. Beyond the constant conjunction of similar objects, and the consequent inference from one to the other, we have no notion of any necessity or connexion. (EHU 64)

The several instances of resembling conjunctions lead us into the notion of power and necessity . . . Necessity then is the effect of this observation and is nothing but an internal impression of the mind. (T 165)

’Tis the constant conjunction of objects, along with the determination of the mind, which constitutes a physical necessity. (T 171)



The word ‘constitutes’ confirms that Hume makes no attempt to point to connexions that lie beyond our conceptual understanding. Accordingly, his two definitions of ‘cause’ exclude all trace of any attempt to give expression to the Bare Thought. Indeed, the connection between what is necessary and what we experience is now so tight that it seems to leave little room for any contingent, indeterminist, causal connections of the sort found in quantum physics, namely, where A causes  $x$  but could have *ceteris paribus* just as easily caused  $y$  or  $z$ <sup>47</sup> (let alone ‘so absurd a Proposition as *that any thing might arise without a Cause*’<sup>48</sup>), a point reaffirmed in his discussion ‘of liberty and necessity’:

... it is pretended that some causes are necessary, some not necessary ... let anyone *define* a cause, without comprehending, as part of the definition, a *necessary connexion* with its effect; and let him show distinctly the origin of the idea, expressed by the definition; and I shall readily give up the whole controversy ... this regular conjunction produces that inference of understanding, which is the only connexion, that we can have any comprehension of. Whoever attempts a definition of cause, exclusive of these circumstances, will be obliged either to employ unintelligible terms or such as are synonymous to the term which he endeavours to define. (EHU 74)<sup>49</sup>

Necessity may be defined in two ways, conformably to the two definitions of *cause*, of which it makes an essential part. It consists either in the constant conjunction of like objects, or in the inference of the understanding from one object to another ... as long as the meaning is understood I hope the word can do no harm. (EHU 75)

Hume’s parallel definitions of the terms ‘cause’ and ‘necessity’ reveal him to take causal claims to describe little more than constant conjunction and human expectation, of which we have an ‘internal impression of the mind’, whose existence is not in any doubt. We must consequently reject the very *idea* of a necessary connection in a robust metaphysical sense, for such an idea (from here onwards capitalised as ‘Necessary Connection’) would be devoid of any meaning whatsoever.

According to Millican, this reveals Hume to be antirealist about Causation (with a capital ‘C’) yet ‘thoroughly realist about objects, events, their resemblances, and their spatial and temporal relationships ... and hence about causation (small ‘c’) as understood in accordance with his first ‘definition of cause’, which includes a notion of necessity (with a small ‘n’).<sup>50</sup> While in general agreement with Millican’s exegesis of Hume’s text, I find his terminology misleading. We should not label Hume an ‘antirealist’, because this suggests (as we have already seen the ‘regularity’ interpreters to have thought) that Hume is denying the existence of something, thereby implying that he takes us to have a meaningful idea of what is being denied



(Millican believes that the implicature can and should be cancelled). If one cannot meaningfully state that there *is* such a thing (Necessary Connection), then, by the same token, it also follows that one cannot meaningfully state that there *isn't* one, or that—for all we know—there *may or may not* be one.

We should also resist labelling Hume a realist about 'small c' causation. While Hume does not doubt its existence, his talk about it is empirical not metaphysical. If he is a realist about causation it is only in the (incredibly weak) sense in which dogs or children are realists about their food.<sup>51</sup> The thought that there is food on the table no more assumes realism than the opposite thought does antirealism (however temporary). To attribute such metaphysics on the basis of empirical remarks is to mistakenly treat a form of expression (e.g., 'there are causes') as if it were a move (an expression) *within* a language game, rather than a constitutive rule.<sup>52</sup>

Given these considerations I think it best we stop thinking of Hume as a sceptical realist, antirealist, quasirealist, or standard realist about either C/causation or N/necessity. On Hume's view as I have been interpreting it, we cannot meaningfully say *anything* about Necessary Connexions, for we have no meaningful idea of such things. Nor is there any meaningful idea of Necessary Connection available in our minds for us to *project* unto the world (we cannot meaningfully talk 'as if' there were Necessary Connections in the world if the very notion of such a thing is meaningless). If we must use philosophical labels, Hume is an empirical realist about necessity, thus tempering both sceptical and metaphysical realism with revisionary semantic strategies which he arguably also employs in relation to ethics,<sup>53</sup> induction,<sup>54</sup> personal identity,<sup>55</sup> and even God.<sup>56</sup>

## VII. BRINGING IT ALL BACK HOME

What of all the references to nature's secret powers and connexions? I end with a brief commentary on two such passages. Here is the first:

[N]ature has kept us at a great distance from all her secrets, and has afforded us only the knowledge of a few superficial qualities of objects; while she conceals from us those powers and principles on which the influence of these objects entirely depends. Our senses inform us of the colour, weight, and consistence of bread; but neither sense nor reason can ever inform us of those qualities which fit it for the nourishment and support of a human body. Sight or feeling conveys an idea of the actual motion of bodies; but as to the wonderful force or power, which would carry on a moving body forever in a continued change of place, and which bodies never lose but by communicating it to others; of this we cannot form the most distant conception. But notwithstanding this ignorance of natural powers and principles, we always presume, when we see like sensible



qualities, that they have like secret powers, and expect that effects, similar to those which we have experienced, will follow. (E 29; cf. T 168)

The most striking thing about this is the footnote which Hume attaches to it:

The word, Power, is here used in a loose and popular sense. The more accurate explication of it would give additional evidence to this argument. See sect. 7 'Of the idea of necessary connexion'.

In other words, 'secret' is here being predicated of a notion ('power' and elsewhere also 'connexion') whose accurate account is going to be given in a later section (until which we have *some loose* idea of power because it consists of regularity as well as necessary connexions, and we already have an impression of that). The section referred to, needless to say, is the one where 'power' and 'necessary connexion' are defined empirically, in relation to a mental impression, culminating in T 165 (quoted in §VI). We should read Hume as presenting a *recapitulation* account of the ideas of cause, power, and necessary connexion: once he has told us what these consist in we may return to the start of his *Treatise* and *Enquiry* and reread them with his empiricist (demystifying) definition in mind.<sup>57</sup>

What is secret, on this interpretation, is a constant conjunction which, were it to be observed, would give rise to an impression of expectation. Hidden causes are part of hidden natural history, in principle observable, even if this is not possible with the naked eye. Hume has in mind, then, not some kind of metaphysical nexus or straightjacket but the sorts of regularities observed by molecular physics.

New Humeans frequently also cite the following passage:

My intention never was to penetrate into the nature of bodies, or to explain the secret causes of their operations. For besides that this belongs not to my present purpose, I am afraid, that such an enterprise is beyond the reach of human understanding, and that we can never pretend to know body otherwise than by those external properties, which discover themselves to the senses. As to those who attempt anything farther, I cannot approve of their ambition, till I see, in some one instance at least, that they have met with success. But at present I content myself with knowing perfectly the manner in which objects affect my senses, and their connections with each other, as far as experience informs me of them. This suffices for the conduct of life; and this also suffices for my philosophy, which pretends only to explain the nature and causes of our perceptions, or impressions and ideas. (T 64)

It may *seem* that while Hume has no interest in 'secret causes' and disapproves of the 'ambition' of those who do, he nonetheless allows that there



are such things. His view of the 'enterprise' of penetrating into them as lying 'beyond the reach of human understanding' might further mislead us into thinking that it is a *metaphysical* one rather than one pertaining to physics or biochemistry. But in the appendix to this passage Hume adds the following note:

As long as we confine our speculations to the *appearances* of objects to our senses, without entering into disquisitions concerning their real nature and operation, we are safe from all difficulties, and, can never be embarrass'd by any question. Thus, if it be ask'd, whether the invisible and intangible distance, interpos'd betwixt two objects, be something or nothing: 'Tis easy to answer, that it is *something*, viz. a property of the objects, which affect the *senses* after such a particular manner. If it be ask'd, whether two objects, having such a distance betwixt them, touch or not: It may be answer'd, that this depends upon the definition of the word, *touch* . . . If we carry our enquiry beyond the appearances of objects to the senses, I am afraid, that most of our conclusions will be full of scepticism and uncertainty. Thus if it be ask'd whether or not the invisible and intangible distance be always full of *body*, or of something that by an improvement of our organs might become visible or tangible, I must acknowledge, that I find no very decisive arguments on either side; tho' I am inclin'd to the contrary opinion, as being more suitable to vulgar and popular notions. If *the Newtonian* philosophy be rightly understood, it will be found to mean no more . . . Nothing is more suitable to that philosophy, than a modest scepticism to a certain degree, and a fair confession of ignorance in subjects, that exceed all human capacity. (T 638–9)

All Hume has in mind when we talks of secret powers, then, is causal properties that cannot be observed with the naked eye, but only through an 'improvement of our organs', or some helpful tool, such as a microscope. His scepticism is so *modest* (EHU 28) that he paradoxically even *names* some of the 'hidden' Newtonian causes:

These ultimate springs and principles are totally shut up from human curiosity and enquiry. Elasticity, gravity, cohesion of parts, communication of motion by impulse; these are probably the ultimate causes and principles we shall ever discover in nature; and we may esteem ourselves sufficiently happy, if, by accurate enquiry and reasoning, we can trace up the particular phenomena to, or near, these general principles. The most perfect philosophy of the natural kind only staves off our ignorance a little longer. . . . (EHU 26)

Clatterbaugh finds this passage 'very strange', deeming that Hume 'seems to both endorse hidden general causes and to deny their value' and that



‘it seems odd to say that ultimate springs are beyond enquiry and then in the next sentence identify some of these ultimate springs’.<sup>58</sup> Clatterbaugh’s worry is not that Hume pleads ignorant of the *underlying* causes of, say, elasticity, but that he seems to deny that one could *ever* know the very ultimate causes of any mechanism when, as a good empiricist, he should think it only a matter of time before these are revealed to us.

There is more than one possible explanation here. Perhaps Hume believed there to be an infinite regress (as suggested by EHU 28), or maybe, as Clatterbaugh suggests, he attempts to embrace Newtonian science while rejecting Newton’s own understanding of causation. My own understanding is that Hume’s pessimism is directly motivated by the more general belief that reason can never justify empirical methods (EHU 28). This famously leads him to endorse a pragmatic Pyrrhonism, according to which it is psychologically impossible to systematically doubt experience.<sup>59</sup> It is no wonder, then, that ‘the most perfect philosophy of the natural kind only staves off our ignorance a little longer’. But whatever the precise details of what Hume takes to be hidden from view here, and why, I hope to have shown that it is not a ‘cause’ in any sense involving necessary connexions that lie beyond anything anticipated by his two definitions. There is truth in all approaches to Hume but, unlike so many of his interpreters, le bon David created new wineskins in which to preserve his new wine.<sup>60</sup>

## NOTES

1. The terms ‘regularity’ and ‘constant conjunction’ are often used interchangeably by commentators. But while Hume makes no explicit distinction between them, he often pairs them in his later work, e.g., EHU 29, 68, 90. Moreover, a close reading of T 87 suggests that ‘regularity’ could refer to the *order* of the contiguity and succession that together constitute constant conjunction. References to both the *Treatise* (T) and the first *Enquiry* (EHU) are to the Selby-Bigge (2nd & 3rd) editions, in standard form (all emphasis in the originals).
2. The passages that follow are quoted in Galen Strawson, “David Hume: Objects and Power”, in *Reading Hume on Human Understanding*, edited by Peter Millican (Oxford: University Press, 2002), 231.
3. Saul Kripke, *Wittgenstein on Rules and Private language* (Oxford: Blackwell, 1982), 67.
4. Roger Woolhouse, *The Empiricists* (Oxford: University Press, 1988), 149–50.
5. Barry Stroud, *The Quest for Reality* (Oxford: Oxford University Press, 2000), 11.
6. Anthony Flew, *David Hume: Philosopher of Modern Science* (Oxford: Blackwell, 1986), 71.
7. J. L. Mackie, *The Cement of the Universe: A Study of Causation* (Oxford: Clarendon Press, 1974), 6; cf. 193ff.
8. G. H. von Wright, *Causality and Determinism* (New York and London: Columbia University Press, 1974), 54.
9. Roger Scruton, *Modern Philosophy: An Introduction and Survey* (London: Sinclair-Stevenson, 1994), 175.



10. For a critical overview of Hume's reasons for revising the definitions first offered in the *Treatise*, see "The Energy of the Cause" in A. Baier, *Death and Character* (Cambridge: Harvard University Press, 2008), 224–36. I sidestep the debate on whether the definitions are intended to be materially equivalent (cf. J. A. Robinson, "Hume's Two Definitions of 'Cause' ", *The Philosophical Quarterly*, Vol. XII, 1965, 129–47), save to say that one way of understanding what he is up to here is that he is offering two different but (in his eyes) equally legitimate uses of the term 'cause'. Paul Russell ingeniously maps these onto his ontology of double existence (*Freedom & Moral Sentiment: Hume's Way of Naturalizing Responsibility*, Oxford: Oxford University Press, 1995, 35–37).
11. Winkler, "The New Hume", 73, emphasis in original.
12. *Ibid.*, 53–54.
13. H. W. Noonan, *Hume on Knowledge* (London: Routledge, 1999), 142.
14. H. E. Allison, *Custom and Reason in Hume* (Oxford: Oxford University Press, 2008), 187.
15. *Custom and Reason in Hume*, 180.
16. *Ibid.*, 181.
17. But see note 21.
18. Allison, *Custom and Reason in Hume*, 372; the interpretation of Malebranche is taken from S. Nadler, "Malebranche on Causation", in *The Cambridge Companion to Malebranche*, edited by S. Nadler (Cambridge: Cambridge University Press, 1990), 112–38. On this reading, Hume only parts company with Malebranche in rejecting his alternative proposal that the only genuine cause was God. For additional possible differences between Hume and Malebranche here, see S. Buckle, *Hume's Enlightenment Tract: The Unity and Purpose of and Enquiry Concerning Human Understanding* (Oxford: Oxford University Press, 2001), 191ff. Buckle allows that the debt to Malebranche's *Search after Truth* is 'so profound' that 'if Hume were a modern academic, he would not escape the charge of plagiarism' (*Hume's Enlightenment Tract*, 191).
19. K. Clatterbaugh, *The Causation Debate in Early Modern Philosophy: 1637–1739* (London: Routledge, 1999), 197–8.
20. For a recent example, see S. Boulter, "Hume on Induction: Genuine Problem or Theology's Trojan Horse?", *Philosophy*, 77 (2002): 299, 67–86, at 70–71. Mackie defends Hume against the accusation that he conflates logic with psychology (*The Cement of the Universe*, 27). To this we might add that such confusions would be incompatible with his allowing that induction may be perfectly justified (EHU 29). For related debates see A. S. Roth, 'Causation', in *The Blackwell Guide to Hume's Treatise*, edited by S. Traiger (Oxford: Blackwell, 2006), 97; N. Goodman, *Fact, Fiction and Forecast* (Cambridge: Harvard University Press, 1954, 4th edition 1983); and. Rupert Read, "The New Antagonists of "the New Hume": On the Relevance of Goodman and Wittgenstein to the New Hume debate", in Read & Richman, *The New Hume Debate*, 167–97.
21. Russell, *Freedom & Moral Sentiment*, 24–42, particularly 35–37. Cf. note 11.
22. For example, the first half of Janet Broughton's argument against a psychological reading simply attacks the view being attributed to Hume, "Hume's Ideas about Necessary Connection", 225ff.
23. A. J. Ayer, *Hume* (Oxford: Oxford University Press, 1980), 64. Cf. A. J. Ayer *Probability and Evidence* (London: Macmillan, 1972), 10–18, quoted with critical commentary in Mackie's *The Cement of the Universe*, 23–24 (fn. 28).



24. S. Blackburn, "Hume and Thick Connections", in *Reading Hume on Human Understanding*, 259–76. If Blackburn is right, then Hume occasionally creates needless puzzlement by conflating the two.
25. For the argument that Hume treats the missing shade of blue as a *type* of exception rather than a single instance, see W. H. Williams, "Is Hume's Shade of Blue a Red Herring?", *Synthese*, 92 (1992): 83–99, at 86.
26. For a captivating history of this footnote throughout the various editions of the *Enquiry*, see Winkler, "Causal Realism and Hume's Revisions of the *Enquiry*" (forthcoming).
27. Cf. J. O Nelson, "Hume's Missing Shade of Blue Re-viewed", *Hume Studies*, Volume XV (1982), 2: 353–64.
28. P. J. E. Kail, *Projection and Realism in Hume's Philosophy* (Oxford: Oxford University Press, 2007, 32).
29. Ibid.
30. Kail, *Projection and Realism*, 33.
31. Ibid., 32.
32. "David Hume: Objects and Power", 237–8, 246–7.
33. Strawson concedes that Hume cannot officially allow that such ideas could have the positive content required for the terms we use to pick them out to mean what they would if we had the impressions in question but argues that there remains a sense in which the constraint still allows him to conceive of such things as secret powers. Cf. *The Secret Connexion—Causation, Realism, and David Hume* (Oxford: Clarendon Press, 1989), 121, 231.
34. *Realism and Projection*, 84.
35. R. Cudworth, *The True Intellectual System of the Universe* (New York: Garland Pub., R. 1978 [1678]), 200.
36. Cf. T. Stoneham, "Berkeley's 'Esse Is Percipi' and Collier's 'Simple' Argument", *History of Philosophy Quarterly* (2006): 211–24, and "The Very Idea of a Dictionary" (unpublished appendix).
37. Ibid.
38. S. Johnson, *A Dictionary of the English Language* (London: J. and P. Knapton, et al., 1755), entry for 'intelligible'.
39. Ibid., entry for 'signification'.
40. EHU 59. Peter Millican, "Hume's Idea of Necessary Connexion: Of What Is It the Idea?", *Hume Society Conference, Iceland* (August): 1–2, points out that 'the thought that there might turn out to be no relevant impression, and hence no idea, is not even raised in the *Enquiry* until the quest is well under way'.
41. *The Cement of the Universe*, 20.
42. Ayer criticises him for this in *Hume*, 68.
43. "Hume's Idea of Necessary Connexion", 3.
44. John P. Wright, *Hume's 'A Treatise of Human Nature'* (Cambridge: Cambridge University Press, 2009), 79.
45. Cf. P. Millican, 'Hume's Idea of Necessary Connexion: Of What Is It the Idea?', *Hume Society Conference, Iceland* (August), §2.
46. *Hume's 'A Treatise of Human Nature'*, 80–81.
47. Whatever point is being made in his discussion of induction, then, is not one that he would have taken quantum physics to affirm. On the contrary, the reasoning process by which it demarcates indeterministic causation is subject to Humean suspicion.
48. "Letter to John Stewart", February 1754, in *The Letters of David Hume*, edited by J. Y. T. Craig (Oxford: Oxford University Press, 1932), vol. 1, 187.
49. In a footnote to this passage Hume adds: ' . . . this constancy forms the very essence of necessity, nor have we any other idea of it'.



50. P. Millican, "Against the 'New Hume'", in *The New Hume*, edited by Read and Richman, 215.
51. I suppose we could distinguish between 'Realism' and 'realism', but see no obvious benefit in the retention of deflated metaphysical concepts.
52. Cf. L. Wittgenstein, *Philosophical Investigations*, 4th edition, translated by G. E. M. Anscombe, P. M. S. Hacker, & J. Schulte (Oxford: Wiley-Blackwell, 2009 [1953]), §§401–2.
53. A. Baier, *A Progress of Sentiments: Reflections on Hume's Treatise* (Cambridge: University Press, 1991) & "Much Obligated", in C. Pigden (ed.), *Hume on Is and Ought* (Basingstoke: Palgrave Macmillan), 2010; C. Sandis, "Hume and the Debate on Motivating Reasons", in *Hume on Motivation and Virtue*, edited by C. Pigden (Basingstoke: Palgrave Macmillan, 2009).
54. P. F. Strawson, *Scepticism and Naturalism: Some Varieties* (London: Methuen & Co., 1985).
55. T. Penelhum, *Themes in Hume: The Self, the Will, Religion* (Oxford: Oxford University Press, 2000), Ch. 2 & 3.
56. R. H. Popkin, "David Hume: His Pyrrhonism and his Critique of Pyrrhonism", *The Philosophical Quarterly*, 1 (1951) 5: 385–407, at 81–84; cf. E 57.
57. The term 'recapitulation' has its intellectual home in a certain strand of theology devised to make sense of the messianic secret in Mark's Gospel. See W. Wrede, *The Messianic Secret*, translated by J. C. G. Grieg (Cambridge: James Clarke & Co., W., 1901/1971).
58. *The Causation Debate in Early Modern Philosophy*, 202.
59. EHU 162; R. H. Popkin, "David Hume: His Pyrrhonism and his Critique of Pyrrhonism", and Penelhum, *Themes in Hume*, vii–ix, 177–203.
60. This chapter was funded by the project *Aspectos Modales del Realismo Materialista*, HUM 2007–61108, MCYT—Spanish Government. It was first presented at the *British Society for the History of Philosophy* conference *Causation: 1500–2000* at the University of York (March 25–7, 2008) and subsequently at the University of Birmingham, *Royal Institute of Philosophy* lecture series (May 6, 2008) and the University of Athens, *Research Seminar in History and Philosophy of Science* (June 15, 2009). Many thanks to the organisers and participants of these events. Thanks also to Keith Allen, Annette Baier, Helen Beebe, Stephen Boulter, Philip Goff, Peter Kail, Dan O'Brien, Peter Millican, Luke Mulhall, Stathis Psillos, Tom Stoneham, Galen Strawson, and Kenneth Winkler for helpful conversations and exchanges. I am particularly grateful to Tom Stoneham for sharing his research on Johnson's dictionary and philosophical texts by Hume's contemporaries.



## 9 Is Causation a Relation?

*Boris Hennig*

Hume thinks that if causation is anything at all it must be a relation between two distinct items. So does almost everyone after Hume. It is debated what kind of relation causation is and among what kinds of items it obtains. For instance, it has been argued that two items stand in a causal relation if there is a law of nature that connects them,<sup>1</sup> or if one of them can be manipulated by manipulating the other,<sup>2</sup> or if some underlying mechanism can be specified that ties them together,<sup>3</sup> or if conserved quantities are transferred from one to the other.<sup>4</sup> As for the related items, some think that they must be two distinct events;<sup>5</sup> others claim that they are rather facts<sup>6</sup> or properties.<sup>7</sup>

In any case, it is assumed that causation is some kind of relation between two distinct items of some sort. This assumption has only rarely been questioned. One good reason for this is that given any causal process, it seems easy to come up with definitions of *R*, *a*, and *b*, such that the relational statement *R(a,b)* is true if and only if the causal process occurs. For instance, if one ball pushes another one, there is a relation between the two balls and their movements. This is why D. H. Mellor, for instance, begins his argument against the assumption that causation is a real relation by distinguishing between relations and real relations. According to Mellor, *R(a,b)* stands for a real relation only if both *a* and *b* exist. Thus he agrees with the general assumption that all causal processes may be represented by statements of the form *R(a,b)*, and only objects that some things that may be represented by *R(a,b)* are not *real* relations.<sup>8</sup> David Lewis does the same in his argument against the claim that causation is a relation. He argues that even though the void is nothing, it has causal properties. Since he also assumes that relations cannot hold among absences, he concludes that causation cannot be a relation.<sup>9</sup> I do not find this argument convincing. It seems to me that ‘better than’, for instance, is a real relation, even though there are many things that are better than nothing. So it cannot be true that real relations cannot hold among absences. Nonetheless, I agree with Lewis’s conclusion: Causation is not a relation.

Causation is not a relation because it must be more than a relation. If it were just a relation, it would be possible to obtain an adequate representation



of a causal process by merely sticking two independent representations of the two related items together. For in general, relations may be identified with n-tuples of items, and this should also be true for causal relations. They should be adequately representable by pairs of items. I maintain that causal processes cannot be adequately represented by pairs of items. There may be causal relations, for which this is true. But causations (i.e., causal processes) are not causal relations, because it is not true for them. It is one thing to be able to split up a process into distinct parts, another to be able to adequately represent the process by simply putting these parts together. I believe that although one may decompose causal processes into parts that stand in causal relations, causal processes are not simply the result of putting these parts together. The causal relation between a cause and its effect is not the same as the causal process of which they are parts.

I will show this by first arguing against what seems to be Hume's argument for the assumption that causation is a relation. Then I will demonstrate that in an example case that Hume discusses, causation cannot be understood as a mere relation between two distinct items. That is, either the two related items are indeed distinct and separable, in which case their causal relation will not be apparent and the items will not be seen as parts of a causal process, or, on the other hand, the two items are represented as causally related, but then their representations are not separable. By Hume's standards, the latter means that there were no two distinct items to begin with, but only one.

Some authors may seem to also deny that causation is a relation. For instance, Norwood R. Hanson claims that the 'causal chain analogy' is misleading.<sup>10</sup> However, since not every relation is a chain, this does not amount to rejecting the idea that causation is a relation. Also, Richard Taylor denies that causation generally is a relation between two events, but he does not deny that it is a relation.<sup>11</sup> Further, David Fair, Wesley Salmon, and Phil Dowe argue that causation is a certain kind of process.<sup>12</sup> If there were a generally accepted distinction between processes and relations such that no relation can be a process, this would imply that causation is not a relation. But Fair, for one, does not draw any such distinction. He has no difficulties writing that causation is 'a physically specifiable relation of energy-momentum flow',<sup>13</sup> and thus thinks that a process such as energy-momentum flow can also be a relation. I do not think so.

Anjan Chakravartty likewise seems to use 'relation' in a way so that it can also refer to causal processes. He writes that although one should conceive of causation 'not as a relation between discrete entities', causal processes still are 'systems of continuously manifesting relations between objects with causal properties and concomitant dispositions'.<sup>14</sup> If by 'continuously manifesting relation' he means something that is a process rather than a relation between distinct items, he might be making the claim that I am going to defend here, except that he sticks to the label 'relation'. However, the other elements of his characterisation of causal processes



are rather obscure to me. For instance, in order to understand this characterisation, one needs to understand what ‘causal properties and concomitant dispositions’ are. According to Chakravartty, causal properties are properties that confer dispositions for behaviour on objects.<sup>15</sup> However, since actualisations of dispositions are conceptually prior to dispositions, one must know in what ways these dispositions are actualised in order to know what kinds of dispositions they are. Presumably, they are actualised by some kind of causal process. This means that in the end, causal properties must be understood to be properties that dispose things to take part in causal processes. If this is so, it is no wonder that Chakravartty can define causal processes in terms of causal properties. Causal processes turn out to be systems of relations among objects that dispose each other to take part in causal processes. If this is what Chakravartty wants to say, it is circular.

## HUME ON CAUSATION

According to Hume, causation is not only a relation, but what he calls a ‘philosophical relation’. A philosophical relation can be any ‘circumstance, in which, even upon the arbitrary union of two ideas in the fancy, we may think proper to compare them’.<sup>16</sup> Philosophical relations can thus also hold among absences. In order for two items to stand in a philosophical relation, they need not both exist, and there need not be a real connection between them. There is only one thing that Hume will not allow: Nothing can stand in a philosophical relation to itself in one and the same respect. Philosophical relations can only hold among distinct items.<sup>17</sup> They are irreflexive. Even the relation that Hume calls ‘identity’ is not actually a relation between a thing and itself, but only a relation between two different stages of the same thing (*Treatise* I,iii,2, p. 74). This is why the relation of identity is common only to all things ‘whose existence has any duration’ (*Treatise* I,i,5, p. 14). Since all transitive and symmetric relations must also be reflexive, this also implies that philosophical relations cannot be transitive and symmetric. If causation is a transitive philosophical relation, it must be asymmetric.

In any case, causation can only be a relation between two distinct items, or two distinct parts, aspects, or moments of one item. For Hume, this has important consequences. He thinks that whenever two items are in any way distinct, they must also be completely separable from each other. He writes that ‘there are not any two impressions which are perfectly inseparable’ (*Treatise* I,i,3, p. 10). This is sometimes referred to as the Separability Principle.<sup>18</sup> It implies that if cause and effect are two items rather than only one, our ideas of them must be distinct, separable, and independent. In this way, the Separability Principle straightforwardly leads from the assumption that causation is a philosophical relation to Hume’s claim that ‘the



effect is totally different from the cause, and consequently can never be discovered in it'.<sup>19</sup>

Hume's assumption that causation is a relation thus implies one of his main claims: that nothing about the effect can be discovered in the cause. As far as I can see, there is only one passage where Hume comes close to defending the assumption that causation is a relation. He writes:

Let us therefore cast our eye on any two objects, which we call cause and effect, and turn them on all sides, in order to find that impression, which produces an idea of such prodigious consequence [as the idea of causation]. At first sight I perceive, that I must not search for it in any of the particular *qualities* of the objects; since, which-ever of these qualities I pitch on, I find some object, that is not possest of it, and yet falls under the denomination of cause or effect. And indeed there is nothing existent, either externally or internally, which is not to be consider'd either as a cause or an effect; tho' 'tis plain there is no one quality, which universally belongs to all beings, and gives them a title to that denomination. The idea, then, of causation must be deriv'd from some *relation* among objects; and that relation we must now endeavour to discover. (*Treatise* I,iii,2, p. 75)

If this is an argument for the claim that causation is a relation, it proceeds as follows. First, it is assumed that causation must be either a quality or a relation. Then it is argued that since we cannot discover a quality that all causes or effects share, causation cannot be a quality. Therefore, Hume concludes, it must be a relation. Every one of these steps is questionable. As for the first, it does not seem to be true that causation can only be a quality or a relation. For instance, causation might also be a process or an event. Further, there are in fact qualities that all empirical objects share, and it is not obviously wrong to suppose that these qualities are somehow related to the fact that all empirical objects are causally related to further empirical objects. There might be ways of construing what Kant says about causality in this way. For now it only matters that Hume's claim, that there are no qualities that all causes share, does not seem obviously true. Finally, Hume infers that causation cannot be a quality from our inability to discover a quality that all causes share. Later, however, he also argues that we cannot discover a relation that holds between all causes and their effects. By parity of reasoning, he should infer from this that causation is also not a relation. However, Hume does not do this. He rather claims that although we cannot discover a causal relation between a single cause and its effect, we can discover such relations by considering not only one but many instances. The same could be said in order to defend the claim that causality is a quality: We may not be able to observe such a quality in a single instance, but we may get to know the causal qualities of things when we get accustomed to dealing with many instances of their kind.



## DROWNING

Hume does not provide a convincing argument that if causation is anything at all it must be a relation. This does not imply that causation is not a relation. In order to show this, I will now consider the following instance of a causal process given by Hume:

The idea of sinking is so closely connected with that of water, and the idea of suffocating with that of sinking, that the mind makes the transition without the assistance of the memory. (*Treatise* I,iii,8, p. 104)

According to Hume, we cannot observe causal connections in one single instance, but given enough experience, we do get to associate effects with their causes. In the present case, we get to associate suffocation with water. Since there is no necessary or conceptual connection between water and suffocation, we cannot discover this connection by merely considering our idea of water. As Hume writes, the first man on earth ‘could not have inferred from the fluidity and transparency of water that it would suffocate him’ (*Enquiry* IV,i, §23, p. 27). We need experience in order to know that water causes suffocation.

The problem with this is that water, as such, does not actually cause suffocation, any more than its concept implies suffocation. It is simply not true that whenever human beings encounter water, they suffocate. This is not even true in the majority of cases. Humans drink water, step into puddles, water their plants, brush their teeth, and so on, without the slightest danger of suffocation. Hume cannot possibly mean that when we get to know the causal properties of water, we come to associate water with suffocation, for it is wrong to generally associate water with suffocation. Someone who thinks that water in itself causes suffocation does not have a reasonable causal belief but a serious problem. Water may be a cause of suffocation, but this does not mean that water usually leads to suffocation. Therefore, when Hume writes that water has ‘suffocated every human creature’ (*Enquiry* VI, §47, p. 57), he cannot actually mean that suffocation is a causal consequence of the mere encounter of humans with water. The cause of suffocation must be something more specific.

Now if we set out to describe the cases in question so that it becomes reasonable to expect suffocation, it turns out that to the extent to which we succeed in giving a description of the encounter with water that justifies the expectation that one will suffocate, we also cease to separate the cause from its effect. As I have said, an encounter with water is not in itself a good reason to expect suffocation. It is one only when one encounters water in a way that leads to a sufficiently severe impairment of respiration. However, an encounter with water that impairs respiration in this way already is an instance of suffocation. That is, the description of an encounter with water that licenses the expectation that one will suffocate is already a description of the beginning of suffocation. There are good reasons to expect suffocation only if a



part of this process is already happening, or if something else is already happening of which suffocation will typically be a later stage.

Hume writes that our idea of water does not imply that we will sink and suffocate, and he is right. However, water does not cause sinking or suffocation any more than our idea of it implies it. There is no reason to generally suppose that water will lead to suffocation on the basis of past experience. There is reason to suppose that a certain kind of encounter with water will lead to suffocation, but when we consider this kind of encounter with water, it turns out that our idea of it does imply that it will lead to suffocation, unless something intervenes. The idea of the kind of encounter with water that causes suffocation is in fact the idea of the beginning stages of suffocation. This means that on the one hand, to the extent that we separate cause and effect so that they are conceptually independent, we fail to see them as causally related. On the other hand, to the extent to which we represent the cause as a cause of an effect, we fail to represent cause and effect as two distinct and separate items.

When we describe an encounter with water as a cause of suffocation, we describe it as the beginning of an instance of drowning, and when we describe an instance of suffocation as caused by an encounter with water, we describe it as the end of an instance of drowning. The kind of encounter with water that leads to suffocation and the instance of suffocation that is caused by this encounter with water are different parts of the same instance of drowning. If there is a relation between cause and effect in this case, it is very much like Hume's relation of identity: it holds between two stages of the same occurrence.

I take it that what has been said about drownings, encounters with water, and suffocation also holds true for other kinds of causal processes. Consider the list of causal verbs that Anscombe gives in "Causality and Determination":

A small selection: *scrape, push, wet, carry, eat, burn, knock over, keep off, squash, make* (e.g. noises, paper boats), *hurt*.<sup>20</sup>

If I am right, one should not think of any of these as a relation between distinct items. If pushing has the same logical structure as drowning, it is wrong to split it up into distinct stages and then say that the pushing is really only a relation between these stages. Rather, the beginning of a causal process may generally be taken to cause the rest of it, and to represent something A as a cause of something else B is to represent A as the beginning of a process of which B typically is a later stage. At the time at which an encounter with water causes suffocation, the drowning of which both the encounter with water and the suffocating are parts is already going on. To separate the encounter with water from the drowning and its later stage, suffocation, is to conceal the fact that it causes suffocation. To bring out this fact is to describe it as a part of an instance of drowning, which leads to suffocation.



Causation is not a relation because, in order to describe a causal process, one must do more than merely refer to two independently described items and add the information that they are causally related. In the case described by Hume, this is not enough because it is not the case that water, as such, causes suffocation. What causes suffocation is something more specific, which is also conceptually related to suffocation. (It is an encounter with water that impairs respiration for a sufficiently long time.)

This can also be seen by considering deviant causal chains. Suppose someone proposes to define drowning as an event that occurs when a living being first gets wet and then ends up dead. This is obviously not a good definition, and one of the main reasons for this is that many cases that are not instances of drowning also satisfy this description. Someone might get stabbed in a shower. This situation is not improved by inserting a further distinct stage and saying, for instance, that drowning is an event that consists of someone getting wet, then suffocating, and then being dead. Someone might get strangled in a shower, and this is still not an instance of drowning. Such deviant causal chains are possible as long as the items in the sequence are not conceptually related. As long as they are described in independent ways and not as parts of the same larger process, that they happen one after the other might as well constitute something other than an instance of drowning.

I have argued that cause and effect are parts of one process, and that to describe them as cause and effect is to describe them as parts of the same. Gerard Heymans has claimed something very similar.<sup>21</sup> He argues that whenever we discover causal connections between two apparently distinct items, we discover that, in fact, these items are ultimately only 'two different sides of a singular fact'.<sup>22</sup> If this is true and if all things in the universe are somehow causally interrelated, it follows that ultimately all things in the universe are only so many different sides of a singular fact.

## CAUSATION AND NECESSITATION

Let me point out some ways in which my argument should not be understood. The problem with Hume's example is not that an encounter with water does not necessitate suffocation. Necessitation is neither sufficient nor necessary for causation. Even in cases where a particular encounter with water is inevitably followed by suffocation, the first need not be the cause of the second. That is, the problem is not that the relation between water and suffocation is contingent—it is that, in itself, water does not at all cause suffocation. Other factors must be present, and to represent these factors is to represent the encounter with water as (the beginning of) an instance of drowning. Still, the kind of encounter with water that causes suffocation need not at all necessitate suffocation. For instance, someone may be drowning but then be rescued, so that she did not drown after all. In such cases, there will be a cause of suffocation without its effect (i.e., the beginning of an instance of drowning without the typical end of drowning).



In an early paper, Bertrand Russell has argued that the notion of a cause should be given up, and his argument might seem to resemble mine.<sup>23</sup> However, it depends on the assumption that I have just rejected: that causes must necessitate their effects. Russell argues roughly as follows. Cause and effect are either (a) contiguous in space and time or (b) not. They cannot be (a) contiguous in space and time because this would imply that at least one of them must be temporally extended. However, Russell argues, if the cause is temporally extended, then only the very last bit of it can be actually causally relevant, and if the effect is temporally extended, only the very first bit of the effect can result from the cause. This argument is of dubious validity, but I will not be concerned with it here. I am interested in the argument that applies to case (b), where cause and effect are supposed not to be spatiotemporal contiguous. In this case, Russell argues, something might intervene so that there is no necessary connection between cause and effect. From this he infers that there is no causal connection between cause and effect, which contradicts the assumption that they are indeed cause and effect. Russell thus assumes that an event cannot be the cause of suffocation unless it necessitates suffocation. I do not share this assumption. I do not argue that an encounter with water cannot be the cause of suffocation unless it is necessary that suffocation follows upon it. I only argue that to represent an encounter with water as the cause of suffocation is to describe it in terms that are conceptually related to suffocation, that is, as the beginning of a drowning. Since drownings may be interrupted, I do not suppose that an encounter with water that constitutes the beginning of a drowning necessarily leads to suffocation.

R. G. Collingwood has put forward an argument that resembles Russell's.<sup>24</sup> He argues that an event A only qualifies as the cause of another event B, in a certain sense of 'cause', if A leads all the way to B. That is, in order to qualify as a cause of suffocation, an encounter with water would have to lead all the way to an instance of suffocation. Before someone actually suffocates, Collingwood argues, the event in question should not be called a cause of suffocation. Again, my argument does not depend on any such assumption. Something might be a cause of suffocation without actually leading to anyone's death, for instance, when someone is drowning but is rescued. Still, that the event in question is a cause of suffocation means that it is the beginning of a drowning.

One might object that an interrupted drowning is not a drowning after all, and if no drowning occurs, nothing can be the beginning of a drowning. So it may seem as though the description of an event as the beginning of a drowning can only be true if the drowning is not going to be interrupted, that is, if someone will indeed suffocate. Therefore, if the cause of suffocation is to be the beginning of a drowning, it would have to necessitate suffocation. However, this objection involves the assumption that a process either is a drowning, and then it must lead to suffocation, or it is not a drowning after all, but only an instance of sinking into water so that one would almost have drowned. If this were true, no one could ever be



saved from drowning. Either the person is drowning and therefore nothing will prevent suffocation, or she is not drowning, and then there would be no reason to rescue her. In order to avoid such consequences, we had better assume that someone may be drowning but then not have drowned after all.<sup>25</sup> If we assume this, there may be beginnings of drownings that are not followed by endings of drownings and more generally, causes will not always necessitate their effects.

Note that I also do not suppose that one can only reasonably expect something to happen when this very thing has already started happening. I have argued that the kinds of encounter with water that license the expectation that someone will suffocate are already beginnings of a process of which suffocation typically is a later stage. But of course one may also expect someone to drown before an actual instance of drowning has begun. For instance, someone might be falling off a cliff, and the circumstances might be such that this person will certainly drown. Note, however, that in this case, we may expect the person to drown only because a larger process is already going on, of which drowning is a later stage. Again, our expectation is justified because what we expect is a later stage of some process that is already taking place.

## CAUSAL AND CONCEPTUAL CONNECTIONS

I have claimed that an encounter with water that entitles us to expect suffocation, and thus the kind of encounter with water that causes suffocation, must already be the beginning of a process of which suffocation is a typical later stage. This implies that cause and effect must be conceptually related, at least in the sense that the description of a cause as a cause must involve a description of its effect.

Many authors believe that causal connections must be contingent. They take for granted that causes cannot logically imply their effects in the same respect in which they cause them. Moritz Schlick writes, for instance, that the amputation of a limb should not be taken to cause the loss of a limb, because it is the loss of a limb by certain means.<sup>26</sup> He seems to assume that because the concept of amputation already involves the loss of a limb, this cannot any longer be a causal consequence of amputation. Likewise, John L. Mackie writes that the movement of a cricketer's bat should not be said to cause a drive because it is a drive.<sup>27</sup> Again, conceptual relations are thought to preclude causal connections.

The idea that causes and effects must not be conceptually related is often attributed to Hume. Anthony Dardis, for one, writes that 'Hume taught us that causal connections are not necessary connections'.<sup>28</sup> However, what Hume taught us can hardly be that necessary connections cannot also be causal ones. Hume starts by assuming that there should be some necessary connection between cause and effect and fails to find any such connection. Then he suggests that even in the absence of a necessary



connection, there may be good reasons to associate causes with effects. This does certainly not imply that if there were a necessary connection, there could be no causal connection. Hume never demonstrates that causal connections cannot be necessary. He only argues that they need not be necessary.

Further, if conceptual connections would preclude causal ones, it would be contradictory to say that the cause of B causes B. For it is a conceptual truth that the cause of B causes B, but by assumption, since the cause of B and B are conceptually related, B cannot actually cause B. This shows that the assumption that causal connections must be contingent is false. Calling the cause of B 'the cause of B' cannot make it so that the cause of B is not any longer the cause of B.

Of course, Dardis does not actually want to claim that once we call the cause of B 'the cause of B', it must cease to be the cause of B. He is concerned with explanatory, not causal, relevance. Obviously, when we are looking for the cause of B, we are not interested in the information that it is 'the cause of B'. We want to know something further about it. To merely say that the cause of B is the cause of B is not to provide a causal explanation. This, however, does not mean that the cause of B is not actually the cause of B. I am here interested in causal, not explanatory, relevance, and although we cannot explain B by merely referring to 'the cause of B', the thing we refer to as 'the cause of B' remains the cause of B.

There is another possible reason for demanding that causal connections preclude conceptual connections. Some accounts of causation start from the basic assumption that something A is a cause of another thing B if B occurs as a consequence of A. If one does not restrict the notion of a consequence so that logical and conceptual consequences are excluded, such accounts may entail that a thing causes everything that it implies. Since A implies A, for instance, it may then seem that everything is a cause of itself. Such consequences can be easily avoided by demanding that conceptual connections do not qualify as causal ones.<sup>29</sup> However, it is important to keep in mind that this restriction is only needed in order to exclude certain cases that do not seem to be cases of causation. If it should turn out that there are instances of causation where causes and effects are inevitably conceptually related, the restriction in question should be given up. For it would then rule out cases that should not be excluded. This means that there is no generally valid principle to the effect that causal connections must be contingent.

Cause and effect are indeed separable in that there need not be a conceptual connection between the description of the cause and the effect. For instance, one may describe an instance of sinking into water so that nothing about suffocation is implied. However, if one does so, one does not describe the sinking in question as a cause of suffocation. Rather, one describes the cause of suffocation as something other than a cause of suffocation. Further, if one describes the sinking as a cause of suffocation, the occurrence of the sinking need not necessitate an occurrence of drowning.



The beginning of a drowning is conceptually related to the typical end of a drowning, but the end of a drowning does not necessarily follow upon its beginning. An instance of drowning may be interrupted and suffocation may be prevented. There is thus no tight and direct logical connection between the cause and the effect. Still, they are conceptually related, because in order to describe them as cause and effect, one must describe them as parts of the same.

Mackie argues that although it is possible to describe causes and effects in a way so that they are conceptually related, this does not imply that causal connections are conceptual connections. 'What matters', he writes, 'is that there are no logically necessary connections between the events themselves, or between any intrinsic descriptions of them, however detailed and complete'.<sup>30</sup> He is right. It is possible to describe a cause and its effect in a way so that there is no logical connection between the two descriptions. (When one does this, one does not describe them *as* cause and effect.) Mackie concludes that therefore, the fact that logically related descriptions can be given does not imply that cause and effect themselves are logically related. I agree and add: by the same token, the fact that two conceptually independent descriptions can be given of cause and effect does not prove that they are in fact conceptually independent. Further, since a description of the cause that is conceptually independent of its effect also must conceal the fact that the cause has this effect, I maintain that the description of a cause as cause cannot be conceptually independent of the effect. It must be conceptually related to the effect, not in the sense of making it necessary, but in the sense of describing the cause as the beginning of something that includes the effect as a typical later stage.

## CONCLUSION

Hume argues that cause and effect must be separable because causation is a philosophical relation, and no philosophical relation can hold between a thing and this very thing itself, in the same respect. Therefore, cause and effect must be distinct things, and since every two distinct things can be separated, they must be conceptually independent. I do not object to the idea that one may give separate and conceptually independent descriptions of causes and their effects. I also do not deny that causes and effects, thus independently described, may be said to stand in causal relations. However, when we conceptually separate cause and effect, we also conceal the fact that the cause is the cause of the effect, and that the effect is the effect of the cause. These concealed facts, I argue, cannot be brought back in by establishing a contingent relation between the two distinct and separate items. Rather, in order to bring out the fact that one thing causes another thing, one must describe both of them as parts of one process, and thus establish a conceptual relation among them. For instance, one must describe an



encounter with water as the beginning of a drowning and suffocation as the end of the same instance of drowning in order to make clear that the first causes the second. When one describes cause and effect in this way, they cease to be seen as independent items. They are only two distinguishable parts of the same.

I conclude that causation is not a relation between two distinct items. It is not a relation because it cannot adequately be described by an expression of the form  $R(a,b)$ , where  $a$  and  $b$  are distinct and separate items. If at all, causal processes could be described by expressions of the form  $R(a,b,c)$ , interpreted as 'a is an early stage of process  $c$ , of which  $b$  is a later stage'. (More will have to be added, because it is not true that every early stage of a process is a cause of every later stage of this process.) Here, the related items are more than two, and more importantly, they are not distinct.<sup>31</sup> I have pointed out that, at best, causation is a relation similar to Hume's relation of identity: a relation between two stages of the same. We have now seen that these two stages can only be kept apart as two conceptually independent and distinct items by abstracting from the fact that one is the cause of the other. The result is not, as Hume has it, that the effect 'can never be discovered' in the cause. The only reason why the effect cannot be discovered in the cause is that the cause is not described as its cause. Hume's trick is that he still uses the word 'cause' for something that is clearly described as something other than a cause, so that he seems to be making a general point about the relation between a cause and its effect. In fact, his point is only that when we consider causes and effects in complete separation from each other, we cannot see them as causes and effects. Contrary to Hume, I infer from this that causation cannot merely be a relation between separate items.

We should therefore admit that besides causal relations among separable items, there are also causal processes, which are something more than mere relations among distinct things or stages. Within a Humean framework one usually asks: What is the difference between causal relations and noncausal ones? We should still ask this question, but there is now also another question to be answered: What is the difference between causal processes and processes that are not causal?

## NOTES

1. Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), Ch. 7.
2. Douglas Gasking, "Causation and Recipes", *Mind*, 64 (1955): 479–87.
3. Stuart Glennan, "Mechanisms and the Nature of Causation", *Erkenntnis*, 44 (1996): 49–71.
4. Phil Dowe, "Causality and Conserved Quantities: A Reply to Salmon", *Philosophy of Science*, 62 (1995) 2: 321–33.
5. Davidson, *Essays on Actions and Events*, Ch. 7.



6. D. H. Mellor, *The Facts of Causation* (London: Routledge, 1995), Ch. 13.
7. Evan Fales, *Causation and Universals* (London: Routledge, 1990), 52.
8. Mellor, *The Facts of Causation*, 156.
9. David Lewis, "Void and Object", in *Causation and Counterfactuals*, edited by John Collins, Ned Hall, L. A. Paul (Cambridge: MIT Press, 2004), 281.
10. Norwood R. Hanson, *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science* (Cambridge: Cambridge University Press, 1952), Ch. 3.
11. Richard Taylor, *Action and Purpose* (Englewood Cliffs, NJ: Prentice Hall, 1966), 16.
12. David Fair, "Causation and the Flow of Energy", *Erkenntnis*, 14 (1979): 219–50; Wesley Salmon, *Causality and Explanation* (Oxford: Oxford University Press, 1998); Dowe, "Causality and Conserved Quantities".
13. "Causation and the Flow of Energy", 220.
14. Anjan Chakravartty, *A Metaphysics for Scientific Realism: Knowing the Unobservable* (Cambridge: Cambridge University Press, 2007), 104, 114.
15. Chakravartty, *A Metaphysics for Scientific Realism*, 108.
16. David Hume, *A Treatise of Human Nature*, edited by L. A. Selby-Bigge and P. H. Nidditch (Oxford: Clarendon Press, 1978), I.i.5, p. 13. Abbreviated as *Treatise*, references are by book, part, and section.
17. Helen Beebe, *Hume on Causation* (London: Routledge, 2006), 31.
18. Don Garrett, *Cognition and Commitment in Hume's Philosophy* (Oxford: Oxford University Press, 1997), Ch. 3.
19. David Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, edited by L. A. Selby-Bigge and P. H. Nidditch (Oxford: Clarendon Press, 1974), IV,i, §25, p. 29. Abbreviated as *Enquiry*, references are by section, part (where applicable), and paragraph.
20. Elizabeth Anscombe, "Causation and Determination", in her *Collected Philosophical Papers, Vol. II* (Oxford: Basil Blackwell, 1981), 137.
21. This has been pointed out to me by Christian Schmidt.
22. Gerard Heymans, "Über Erklärungshypothesen und Erklären überhaupt", *Annalen der Naturphilosophie*, 1 (1902): 473–85, quotation from 483.
23. Bertrand Russell, "On the Notion of Cause", *Proceedings of the Aristotelian Society*, N.S. 13 (1912–1913): 1–26, see particularly 5–8.
24. R. G. Collingwood, *An Essay on Metaphysics* (Oxford: Oxford University Press, 1998), 314–15.
25. Cf. Sebastian Rödl, *Kategorien des Zeitlichen. Eine Untersuchung zu den Formen des endlichen Verstandes* (Frankfurt: Suhrkamp, 2005), Ch. VI.
26. Moritz Schlick, *Philosophical Papers, Vol. II* (1925–1939) (Dordrecht, Netherlands: Reidel, 1979), 244.
27. J. L. Mackie, *The Cement of the Universe: A Study of Causation* (Oxford: Clarendon Press, 1974), 288.
28. Anthony Dardis, "Sunburn, Independence Conditions of Causal Relevance", *Philosophy and Phenomenological Research*, 53 (1993) 3: 577–98, quotation from 577.
29. Cf. Ernest Le Pore and Barry Loewer, "Mind Matters", *The Journal of Philosophy*, 84 (1987): 630–42; see particularly 635.
30. Mackie, *The Cement of the Universe*, 17.
31. One may define a binary relation  $R'(a,b) := R(a,b,c)$ . However, this does not show that causation is a relation between two distinct items, since the definition of  $R'(a,b)$  still presupposes the notion of a causal process  $c$  of which  $a$  and  $b$  are parts. Here, causation is not a relation but rather one of the related items, namely  $c$ .



# 10 Kant on Causal Knowledge

## Causality, Mechanism, and Reflective Judgment

*Angela Breitenbach*

Kant is well-known for his transcendental conception of causality. In the *Critique of Pure Reason*, he famously presents the causal law as an a priori principle of human understanding rather than an empirically discoverable fact about the world. According to this principle, as Kant argues in the “Second Analogy of Experience”, every change in nature has a natural cause.<sup>1</sup> We can thus know a priori that relations of cause and effect thoroughly determine all events that occur in the world. Discussions of Kant’s conception of causality usually focus on this transcendental principle.<sup>2</sup> In other parts of his writings, however, Kant’s discussion of the possibility of attaining causal knowledge of the world takes a rather different focus. In particular, in the second part of the *Critique of Judgment*, the *Critique of Teleological Judgment*, Kant is concerned not with causality as the transcendental conditions of experience in general but with the possibility of causally explaining concrete parts of nature and, more specifically, corporeal nature. Kant phrases this discussion in terms of the mechanical explicability of the natural world, where the mechanism of nature, as he tells us, is the determination of nature ‘according to the laws of causality’.<sup>3</sup> Kant presents his account of the mechanism of nature in the context of his theory of living beings. Organisms, he suggests, raise a difficulty for any mechanistic account of the world: they do not seem to be amenable to mechanical explanation. Kant concludes in the *Critique of Teleological Judgment* that we cannot know, but can only assume, that nature is determined mechanically. The principle according to which material nature is thoroughly determined by merely mechanical laws, Kant claims, is a purely regulative and subjective maxim that tells us something about how we must approach nature rather than about the objective character of nature itself.

The combination of these two claims, one developed in the *Critique of Pure Reason*, the other in the *Critique of Judgment*, may seem problematic. For how can we know a priori and with apodictic certainty that every



change in nature has a natural cause, but have no way of knowing and can merely assume that material nature is thoroughly determined by the laws of causality? A first response to these questions may be that between writing the *Critique of Pure Reason* and the *Critique of Judgment* Kant changed his mind and gave up the causal principle as a necessary condition of possible experience.<sup>4</sup> According to this reading, in the third *Critique* we can no longer assume that all of nature is determined causally but must allow that certain parts of nature, notably living beings, fall outside the causal framework. Despite the difficulty of combining Kant's various statements on natural causality, however, the *Critique of Judgment* does not seem to give any further evidence for this reading.<sup>5</sup> Thus, even in the third *Critique*, Kant still seems to allow for the transcendental principle of causality introduced in the *Critique of Pure Reason* as necessary for conceiving of events in nature.<sup>6</sup> How, then, can we make sense of Kant's apparently conflicting claims about causality and the mechanism of nature as forming part of a coherent conception of what it means to have causal knowledge of the world?

My aim in this chapter is to propose an answer to this question that pays particular attention to Kant's account of the study of mechanical causes as a tentative, regulatively guided and reflective activity. I argue that rather than presenting a tension, there is an important continuity between Kant's principle of causality and his regulative principle of mechanism. Kant's discussion of living beings, in the *Critique of Teleological Judgment*, does not give Kant reason to alter, retrospectively, his conception of natural causality. Rather, the apparent anomaly of living nature only brings to the fore, perhaps more pointedly than that of inanimate nature, that our causal knowledge of the world depends not only on the determination of experience by the principle of causality but also on an essentially reflective process in our search for causal explanations.

I develop my argument in five steps. I begin, in Section I, by giving a brief account of the principle of causality, introduced in the *Critique of Pure Reason*, and the maxim of mechanism, presented in the *Critique of Teleological Judgment*. In Section II, I consider two competing tendencies in the literature in understanding the compatibility and relatedness of these two principles. According to the first, the concept of mechanism presents an interpretation of the concept of causality that has no objective ground in the nature of things but is subjectively necessary for beings with a kind of understanding as our own. This interpretation accounts for the fact that only the mechanical but not the causal principle is purely regulative. The reading faces the difficulty, however, that it leaves no room for the possibility of genuine mechanical explanation. It thus conflicts with Kant's conviction that science essentially relies on knowledge of mechanical causes. In order to rescue the possibility of scientific knowledge, the concept of mechanism would have to have objective validity. This, it is proposed according to a second reading, is established by Kant in the *Metaphysical Foundations*



of *Natural Science*. The mechanistic interpretation of causality, he argues there, is necessitated by the objective character of material nature. And yet, while this second reading allows for the possibility of genuine mechanical explanations, it does not account for the regulative status of the maxim of mechanism.

The difficulty, then, lies in making sense of the mechanistic maxim as a purely regulative principle that nevertheless enables us to make determining statements about mechanical causes. As an answer to this difficulty, I develop an alternative reading of the maxim and its relation to the principle of causality. In Section III, I argue that Kant introduces the maxim as a specific interpretation of the regulative principle of the systematic unity of nature. While the mechanistic concept of causality is objectively necessary for the realm of material nature, I suggest in Section IV, the maxim of mechanism is nevertheless purely regulative. It aims at a unified conception of nature as systematically ordered by particular mechanical laws, a conception which remains ultimately unattainable. As a regulative principle, the maxim of mechanism nevertheless provides us with a guide to our search for genuine mechanical explanations. By instructing us to think about nature as a whole as determined by mechanical laws, it leads us to apply the concept of mechanical causation to experience. On Kant's account, knowledge of particular causes in the material world, I conclude in Section V, is thus essentially dependent not only on the transcendental principle of causality but also on the regulative maxim of mechanism. Focussing on Kant's regulative account of the mechanism of nature, I suggest in this chapter, will help us get a better sense of Kant's multilayered account of causal knowledge.

## I. THE PRINCIPLE OF CAUSALITY AND THE MAXIM OF MECHANISM

In the "Second Analogy of Experience", Kant argues that we can know a priori that '[a]ll alterations take place in conformity with the law of the connection of cause and effect'.<sup>7</sup> In his proof of this principle, Kant claims that the concept of causality is necessary for distinguishing an objective sequence of states in the world from the merely subjective sequence of perceptions. Thus, Kant points out that in any experience, whether it be the experience of an event or that of a stationary object, we are confronted with a succession of perceptions. In the case of an event, as opposed to that of an object, however, we conceive of this succession as *necessarily ordered*. To employ Kant's well-known example, when we observe a house we can imagine that the successive perceptions of, for instance, the walls, the windows, and the roof could have been ordered differently. The order of perceptions depends on us, on the way we move around the house and turn our head, for example. When we experience a ship moving down the river,



by contrast, our perceptions of the ship at various points on the river could not have had a different order. In observing the ship move with the current, we could not have seen the ship first downstream and then upstream. We thus take the event, but not the stationary object, to consist of a determinate succession of different states in the world, an objective succession to be distinguished from the merely subjective order of perceptions.

Insofar as the objective order of processes in the world ‘remains undetermined through mere perception,’<sup>8</sup> Kant concludes that it is determined by

the understanding . . . ; and in this case it is the concept of the *relation of cause and effect*, of which the former determines the latter in time, as the consequence, and not as something that could precede solely in the imagination.<sup>9</sup>

Only if we can conceive of the states of an event as necessarily ordered as causes and effects, Kant thus argues, can we regard these states as forming part of an objective process in time. The principle of the relation of cause and effect is an a priori law of the understanding that makes the experience of objective succession possible. It is a constitutive principle insofar as something can be the object of determinate temporal experience only if it falls under the law of causality. As a consequence, Kant denies that we could ever experience supernatural causation or ‘creation’.<sup>10</sup> For something that was not itself caused by anything in nature could not be ordered in relation to other experiences in time and would fall outside the spatiotemporal framework of experience. The principle of causality is thus valid, Kant claims, without exception for all of nature:

It is a universal law of the very possibility of all experience that everything which happens has a cause. Hence the causality of the cause, which *itself happens* or comes to be, must itself in turn have a cause; and thus the entire field of experience, however far it may extend, is transformed into a sum-total of mere nature.<sup>11</sup>

According to Kant, it is thus an a priori and constitutive law that every change in nature has a natural cause. The principle of causality is a universal law of *nature as such*, that is, of all that which can, in principle, be experienced.

So far, this is Kant’s well-known story about the a priori principle of causality. In the *Critique of Teleological Judgment*, however, Kant seems to complicate matters by introducing a further principle of mechanical causation. There, Kant is concerned not with the general concept of nature of the *Critique of Pure Reason*, but with ‘particular experiences’ of material, or corporeal, nature and, more specifically, of living beings as opposed to nonliving things.<sup>12</sup> Thus, Kant argues that we experience organisms, but not inanimate nature, as characterised by a specific organisation. Plants and



animals are distinguished by a particular kind of arrangement of their parts within the whole and by a reciprocal interdependence between these parts. To take Kant's own example again, if we consider 'the structure of a bird, the hollowness of its bones, the placement of the wings for movement and of its tail for steering etc.' we seem to think of the parts of the bird as determined by their function within the organism as a whole.<sup>13</sup> Our understanding of the eye as an organ of the bird, for example, is determined by its function for the working of the whole organism. We conceive of the eye as that organ which enables the bird to see. More generally, the existence and form of the individual parts of an organism seem to be purposive for the existence and survival of the organism as a whole. And insofar as each individual organ is, in turn, dependent on the other organs for its own working, the living being as a whole seems to sustain and bring about itself. The organism thus seems to display not only a particular organisation of the parts within the whole, but also a capacity of the whole for self-organisation.

Crucial to this discussion of living beings is Kant's claim about their mechanistic inexplicability. By reference to purely mechanical laws, Kant suggests, we could not make any sense of the specific purposiveness that we seem to experience in living creatures. Instead, Kant goes on to argue, experiences of organisms 'bring reason into play in order to conduct the judging of corporeal nature and its laws in accordance with a special principle'.<sup>14</sup> Thus, we can only make sense of the particular character of an organism, according to Kant, by means of an idea of reason, that is, the idea of purposiveness. We can think of the special character of organisms in terms of the end-directedness of our own rational activity. This does not mean, as Kant takes pains to point out, that we can actually have any knowledge of such purposive directedness in nature itself. Insofar as purposes are essentially connected with the intentionality of an agent who sets something as a purpose, we cannot discover that nonrational nature is in fact purposive. And yet, Kant argues, we nevertheless consider organisms *as if* they were purposively organised and end-directed. Instead of making determinate statements about the presence or absence of final causes in nature we can thus make assertions only about our reflections on the apparent teleological organisation and directedness of living creatures.

These considerations about organisms lead Kant, in the "Dialectic of Teleological Judgment", to raise the question of the mechanical determination of nature.<sup>15</sup> In particular, he argues that in reflecting about material nature we follow two competing principles. On the one hand, we assume that '[a]ll generation of material things and their forms must be judged as possible in accordance with merely mechanical laws'.<sup>16</sup> On the other hand, by contrast, we follow the maxim that '[s]ome products of material nature cannot be judged as possible according to merely mechanical laws (judging them requires an entirely different law of causality, namely that of final causes).'<sup>17</sup> It is the experience of living creatures, it thus seems, that presents an obstacle to considering all parts of corporeal nature as mechanically determined.



In the following, I shall not be concerned with an interpretation of the conflict that these two principles raise, let alone attempt a solution to it.<sup>18</sup> Central for present purposes is rather the particular status that Kant ascribes to these principles. The principle of the mechanical determination of nature is introduced as one of two regulative maxims that we employ, according to Kant, in our judgments about the material world. While these principles make no determining claim about nature itself, they present regulative maxims about the way we think about nature. The flipside of this is the denial of the possibility of knowing that all corporeal nature stands under mechanical laws. We cannot know but can only reflect about nature, Kant thus argues, as if it were completely mechanically determined.

## II. REGULATIVE AND CONSTITUTIVE PRINCIPLES OF THE MECHANISM OF NATURE

If the maxim of mechanism were supposed to be identical with the principle of causality, then the regulative status of the former would stand in conflict with the constitutive status of the latter. It is thus natural to conclude that there must be an important difference between the two principles. In particular, it has been argued that there is a difference between what Kant understands by ‘causality’ in the *Critique of Pure Reason* and what he refers to under the title of ‘mechanism’ in the *Critique of Judgment*. One of the most influential statements of such a differentiated account of mechanism in the third *Critique* is that proposed by Peter McLaughlin.<sup>19</sup> According to McLaughlin, Kant’s concept of mechanism differs from that of causality by its reference to the relation between material parts and their combination as a whole. To explain a natural object mechanically, McLaughlin argues, is to explain the way in which the parts determine the object as a whole. Thus, ‘a mechanical explanation means the reduction of a whole to the properties (faculties and forces) which the parts have “on their own,” that is, independently of the whole’.<sup>20</sup> This determination of the whole by means of its parts, as McLaughlin puts it, gives order to ‘an *inclusion in space*’.<sup>21</sup> It is therefore not analytically entailed by the concept of causality as such which orders ‘a *sequence in time*’.<sup>22</sup> The reason why we nevertheless interpret causality mechanistically, according to McLaughlin, is explicable by reference to the particular reductionist character of the human understanding. Due to the peculiar nature of our understanding, we simply *cannot but* interpret causality in this way: ‘We cannot regard a “real whole” as the cause of the properties of the parts but only as the effect of these properties.’<sup>23</sup> The peculiar mechanistic character of the human understanding, moreover, is a purely subjective fact, incapable of further justification. It is a fact about us rather than about the world that accounts for the identification of relations in time with those in space.



On McLaughlin's reading, the maxim of mechanism is a specific interpretation of the principle of causality that is only subjectively necessary for the human understanding. By spelling out the way in which creatures with a type of understanding as our own have to approach nature, the principle of mechanism has a purely regulative rather than constitutive status. McLaughlin's interpretation has the advantage of explaining both why there seems to be an important connection between causality and mechanism and why, at the same time, only the causal, but not the mechanical, principle is objectively justified and hence constitutive of nature.

And yet, this reading also raises an important difficulty. For it does not seem to provide us with any possibility of ever attaining determinate mechanical explanations of nature.<sup>24</sup> If the very concept of a mechanical cause is not objectively valid, then there is no hope that mechanical explanations will ever yield objective knowledge. In thinking about nature by means of a concept that has merely subjective status, mechanical explanations would themselves turn out to make a claim about how we are to reflect about nature rather than about the objective character of nature itself. This result is crucially at odds, however, with Kant's explicit statement that we can only explain the particular character of nature mechanically, that is, by reference to efficient causes, or the 'nexus effectivus'.<sup>25</sup> Thus, Kant claims, it 'is of infinite importance to reason not to let the mechanism of nature in its productions be dropped out of sight and be bypassed in its explanations; for without this no insight into the nature of things can be attained'.<sup>26</sup> If mechanical explanations are the only way of achieving scientific insight, then they must not rest on purely subjective considerations, or else scientific knowledge itself will be impossible. McLaughlin's interpretation of the status of mechanical causation thus seems to raise serious problems for the possibility of mechanical explanation in science.<sup>27</sup>

How, then, can we avoid this difficulty? Although, in the third *Critique*, Kant does not spell out his justification for interpreting the causality of nature in mechanistic terms, I believe that we can find such a justification in other parts of his writings. In particular, we can find it in his argument that within the context of *material* nature causal relations have to be interpreted as *external* causal relations. Kant develops this account in the *Metaphysical Foundations of Natural Science*. There, he claims that given the transcendental principle of causality and the empirical concept of matter, we can know a priori that 'every change in matter has an external cause'.<sup>28</sup> Such external causes, he argues, must be understood in terms of the interactions between parts of matter by means of their forces of attraction and repulsion. For only by means of these 'original moving forces', Kant claims, can material objects have an extension in space and impart motion to other parts of matter.<sup>29</sup> In the context of corporeal nature, the principle of causality must therefore be interpreted as a principle of external, that is, mechanical influence. If one thus attempts to explain change in a material object, one has to refer to the external



influence of one material part on another. As Hannah Ginsborg puts it, to explain an object mechanically is to account for it ‘in terms of the fundamental powers of matter as such’.<sup>30</sup> This, I believe, accounts for the reductionism of mechanical explanations that McLaughlin detects in the *Critique of Teleological Judgment*. For, to explain change in a material object, one must refer either to the influence of another material object acting on the first, or to the interactions between simpler material parts of the object standing in external causal relations to one another. Since, in his discussion of the contrast between the explicability of living and nonliving nature in the third *Critique*, Kant is particularly interested in the explanation of the generation and inner functioning of organisms, he is specifically concerned with the second case of forces acting between the simpler material components of a complex material whole. For in order to explain the generation and functioning of such beings, we must refer to the external influence that its material parts exert on one another. We thus identify the cause of such change in time with the external influence of parts of matter acting on one another in space. Consequently, the identification of the relationship between causes and their effects with the relationship between material parts and the wholes that they make up is not a merely subjective fact about the peculiar nature of the human understanding. Rather, it is required by Kant’s conception of mechanical causation proposed in the *Metaphysical Foundations*. Thus, Kant argues,

[i]f we now consider a material whole, as far as its form is concerned, as a product of the parts and of their forces and capacities to combine by themselves (including in our consideration other materials which these parts add to one another), we represent a mechanical kind of generation of it [i.e., the material whole].<sup>31</sup>

The fact that Kant here characterises the mechanical explanation of a material object by reference to the relationship of parts and whole does not entail, contra McLaughlin, that the mechanism of nature is the purely subjectively motivated identification of relations in time with those in space. It shows rather, I suggest, that the causality of material nature must be understood as external causality, that is, as the causal interaction between parts of matter by means of their forces of attraction and repulsion. Empirical mechanical laws can thus be understood as a specific version of causal laws, that is, as laws of external causality. It is because of this that, as Kant argues, ‘a real whole in nature is to be regarded only as the effect of the concurrent moving forces of the parts’.<sup>32</sup>

According to the *Metaphysical Foundations*, we can thus know that material nature is ultimately determined by external, or mechanical, causation. And yet, if we can know a priori that all of material nature is subject to mechanical causes, do we really need to reflect about nature by reference to a further *regulative* maxim of mechanism? By turning to the



*Metaphysical Foundations* in support of the mechanical maxim, would we not render that maxim itself constitutive?<sup>33</sup>

The two contrasted interpretations of the maxim of mechanism thus present us with a dilemma. Either we read the mechanistic principle as a merely subjectively necessary interpretation of the principle of causality, thus conflicting with Kant's explicit claim of the necessity of mechanical explanation in science; or we understand the mechanistic interpretation of the causal principle as objectively necessitated by the character of material nature, thus presenting a challenge to the regulative status of the maxim of mechanism. The question, then, is whether we can make sense of the maxim of mechanism as a purely regulative principle, on the one hand, that nevertheless allows for the possibility of genuine mechanical explanations of nature, on the other.

### III. REFLECTIVE JUDGMENT AND THE UNITY OF NATURE

In order to see how Kant can argue both for the objective necessity of mechanical explanation in science and for the impossibility of knowing that all material nature is determined by mechanical laws, it is instructive to pay attention to the way in which Kant introduces the maxim of mechanism in the "Dialectic of Teleological Judgment". Kant begins, in §70, by telling the reader that our understanding of nature as 'the sum of the objects of the outer senses' is grounded in two types of law.<sup>34</sup> The first type is that of 'the universal laws of material nature in general', the second that of 'the particular laws that can only be made known to us by experience.'<sup>35</sup> By the universal laws of material nature in general, Kant seems to refer to those laws which, according to the *Metaphysical Foundations*, can be known a priori to hold for any part of material nature. On their own, however, these general laws are insufficient for a full account of particular experiences of corporeal nature. It is thus an empirical task to search for more specific laws under which we can subsume particular experiences. Kant argues that in so doing, however, we require a 'principle' or 'guideline' according to which we can investigate nature.<sup>36</sup> Moreover, it is the maxim of mechanism, as Kant subsequently introduces it, that presents such a guideline for our empirical research into the particular laws of nature.

At the beginning of §70, Kant only gives us the outline of an argument for these statements. In particular, he here presupposes the claim, developed first in the *Critique of Pure Reason* and subsequently in the introduction to the *Critique of Judgment*, that we must assume a principle of unity as a regulative principle for the study of nature. In order to understand the status of the maxim of mechanism as a regulative guide to our empirical enquiries, we thus need to take into account Kant's previous arguments. This, as I argue in the next section, will elucidate why, after defending a constitutive principle of causality in the *Critique of Pure Reason* and a



mechanical interpretation of this law in the *Metaphysical Foundations*, Kant introduces a purely regulative principle of mechanism in the *Critique of Judgment*.

Kant first discusses the principle of the unity of nature in the “Appendix to the Transcendental Dialectic” of the *Critique of Pure Reason*.<sup>37</sup> Here, Kant raises the worry that although we may have a priori knowledge of the transcendental laws of nature, these transcendental principles leave our empirical knowledge of the specific laws of nature underdetermined. In the case of the a priori principle of causality, this means two things. First, insofar as the principle provides no secure guarantee for our supposed knowledge of causal laws, future observations and experiments may always prove false what we believe to be the causal laws that determine a particular event. We may thus always be mistaken about particular causal explanations. Second, Kant raises an even more general difficulty about the very possibility of discovering causal regularities. Even though the a priori principle of causality guarantees that every event has some cause, Kant claims that it does not ensure that we will empirically discover any evidence of such a cause. For the natural processes that we experience could be so irregular that we would never have evidence for any causal regularities. Thus, Kant suggests that

[i]f among the appearances which present themselves to us there were so great a variety, I do not want to say according to the form (for in that respect they might resemble one another), but according to the content, that is, according to the manifoldness of the existing entities, that even the acutest human understanding could not by comparison of them detect the slightest similarity (*a case which is quite conceivable*), then the logical law of genera would have no sort of standing, and there would not even be a concept of genus, or indeed any other universal concept.<sup>38</sup>

Similarly, in the published introduction of the *Critique of Judgment*, Kant argues that

it may certainly be thought that . . . it would be impossible for our understanding to discover in [ . . . nature] an order that we can grasp, to divide its products into genera and species in order to use the principles for the explanation and the understanding of one for the explanation and comprehension of the other as well.<sup>39</sup>

According to Kant, it is thus conceivable that we fail to recognise any kind of similarity between different objects of experience. This would mean that we could not classify kinds of experience into genera and species and, hence, that we would have no evidence for any empirical laws.<sup>40</sup> No empirical evidence of causal regularities would thus be available. Even if we can know a priori



that every event in nature has a cause, this knowledge would leave open whether we can find out about the particular causes of any given event.<sup>41</sup>

And yet, although it may thus be conceivable that we fail to know of any particular causal laws, it is nevertheless a practical impossibility to go about our business without assuming a great number of regularities to hold in nature. How else could we expect that having lunch will satisfy our hunger, that taking the train will get us to where we want to go, or that we will wake up well rested after a good night's sleep? Kant thus argues that although we may not know a priori what regularities hold in nature, or whether we can find any such regularities at all, we nevertheless have to approach nature *as if* its causal processes were unified by particular laws. In the "Appendix", Kant thus argues that the idea of the systematic unity of nature is a regulative idea of reason. The faculty of reason is that capacity by which we unify individual cognitions under higher principles and strive to combine all individual cognitions into a systematic whole. In order to make sense of our experiences as coherently connected, we have to assume that nature is unified. In particular, in our attempt to make sense of our experiences of nature as forming a coherent whole, we must presuppose the possibility that nature represents a unity systematically connected by necessary laws. In this sense, the unity of nature is only a '*projected* unity'.<sup>42</sup> The idea of unity, Kant argues, is *read into* our experience of nature and thus presents an idea that does not constitute, but regulates, our experience of nature.

In the two introductions to the *Critique of Judgment*, Kant develops this idea of the unity of nature further. The principle of unity is now presented as a necessary principle of the reflective function of judgment.<sup>43</sup> What Kant here entitles 'reflective judgment' is the capacity to search for general concepts or principles that subsume given experiences. Reflective judgment is distinguished from determining judgment insofar as the former *reflects* about experiences in order to discover concepts and laws that may subsume such experiences, whereas the latter *determines* particular experiences by applying *given* principles to them.<sup>44</sup> The a priori principle of causality, for instance, is given by the understanding a priori and can therefore be applied to particular perceptions by means of the determining function of judgment. The empirical laws of nature, by contrast, are not given to judgment but must be discovered empirically. In order to have any reason for hope that this reflective activity will be successful, however, we need to assume that nature is ordered according to general laws:

judgment must thus assume it as an a priori principle for its own use that what is contingent for human insight in the particular (empirical) laws of nature nevertheless contains a lawful unity, not fathomable by us but still thinkable, in the combination of its manifold into one experience possible in itself.<sup>45</sup>



On the one hand, Kant thus argues that, given what we can know about the world a priori, the validity of any set of empirical laws would seem contingent and altogether arbitrary. On the other hand, Kant also claims that in order to regard our empirical experiences as systematically connected, we have to assume that nature is unified according to necessary laws. The principle of judgment according to which we regard nature as a systematic and lawful unity is thus a merely ‘subjectively necessary transcendental pre-supposition’.<sup>46</sup> Although it does not make any claim about nature itself, it is an assumption necessary for making empirical judgments about nature. It is a principle, Kant argues, required if the faculty of judgment ‘is to have any hope of an interconnected experiential cognition in accordance with a thoroughgoing lawfulness of nature or of its unity in accordance with empirical laws’.<sup>47</sup>

In order to make sense of our search for causal explanations, we must therefore reflect about nature as if it were systematically ordered by more specific natural laws. So far, this investigation does not say anything about the regulative maxim of mechanism. It can now be shown, however, that there is an important continuity between Kant’s arguments for the regulative principle of the purposive unity of nature and his introduction of the maxim of mechanism. For, after arguing that we must assume a regulative principle of reflective judgment according to which nature is systematically ordered by empirical laws, Kant goes on to claim that we must postulate a *more specific* maxim of reflective judgment according to which nature is systematically ordered by *mechanical* laws. What, then, is Kant’s argument for this further claim? In what sense is the mechanistic maxim a necessary assumption of reflective judgment?

#### IV. THE MAXIM OF MECHANISM AND MECHANICAL EXPLANATIONS OF NATURE

In the “Dialectic of Teleological Judgment”, Kant argues that the mechanistic maxim ‘is provided to [ . . . judgment] by the mere understanding a priori’.<sup>48</sup> Given Kant’s characterisation of reflective judgment in the third *Critique*, this statement may seem puzzling. As we have seen, our attempt to discover the empirical laws of nature is characterised as a reflective procedure. We cannot simply apply given laws to particular experiences but must reflect about nature in order first to discover such laws. When Kant introduces the maxim of mechanism as a principle of judgment, yet asserts that this maxim is provided to judgment by the understanding, he cannot therefore mean that the maxim is given in the same sense in which the principle of causality is made available to judgment in its determining function. Rather, in light of what I have argued in Sections 2 and 3, the maxim of mechanism can be provided by the understanding only in an indirect sense.



Thus, Section 2 has shown that the identification of causality with mechanical causality is justified not because of a subjective fact about human understanding, but because of what we can know a priori about material nature. The principle that ‘every change in matter has an external cause’ is a constitutive principle making a determinate claim about corporeal nature.<sup>49</sup> Moreover, Section 3 has set out that the transcendental principle of causality leaves underdetermined the particular nature of specific causal laws. The crucial point to note is now that, just as the causal principle leaves open the nature of particular causal laws, so does the law of mechanical causation spelt out in the *Metaphysical Foundations*. While, according to this law of mechanical causation, we can know a priori that every change in material nature has an external cause, we can only find out about the particular causes of any specific change empirically. In order to explain particular experiences of material nature, we therefore have to investigate how the external interactions of parts of matter can causally affect a change in material objects. And, in order to make sense of this search, we must assume that nature is systematically ordered by particular mechanical laws that can, indeed, be discovered by us. In our search for such particular mechanical causes that affect changes within material objects, we thus have to presuppose the maxim of mechanism: ‘[a]ll generation of material things and their forms must be judged as possible in accordance with merely mechanical laws’.<sup>50</sup>

The maxim of mechanism thus goes beyond the constitutive principle of external causality, spelt out in the *Metaphysical Foundations*. It presupposes that all of material nature stands under particular mechanical laws. It is the relatedness, yet also the crucial difference, of these two principles that can now explain Kant’s claim that the maxim of mechanism is, on the one hand, ‘provided’ to the faculty of reflective judgment ‘by the mere understanding a priori’ and yet, on the other hand, represents a principle of reflective judgment.<sup>51</sup> For, insofar as it is a principle of the understanding that every event in material nature has a mechanical cause, the understanding provides, or makes available, to our reflective faculty of judgment a mechanistic interpretation of the principle of the unity of nature. In our search for mechanical explanations we interpret the regulative principle of the lawful unity of nature in mechanistic terms, thus yielding the maxim of mechanism.

The proposed reading can, I suggest, avoid the difficulties facing the two competing interpretations of the maxim of mechanism contrasted in Section 2. First, although the concept of the mechanism of nature has constitutive status, the maxim itself is purely regulative. Insofar as it is a specific interpretation of the regulative principle of the unity of nature, it provides a guide for our reflections about the causal structure of nature. This is why the maxim makes a claim about how we are to approach nature rather than about the objective character of nature itself. The maxim is not regulative because it asks us to make use of a subjective concept of mechanism but



rather because it is a specification of the more general regulative principle of the lawful unity of nature. The maxim of mechanism has regulative status, in other words, because it expresses a specific interpretation of the subjective principle that nature as a whole is systematically organised according to laws cognisable by the human understanding.

This reading implies that the regulative status of the maxim of mechanism cannot simply be explained by reference to the apparent teleological character of living beings that Kant discusses in the third *Critique*. It is not just the experience of organisms as such which makes it necessary, for Kant, to relegate the mechanical principle to a merely regulative status. Even if living nature could, in other words, be explained according to mechanical laws, the maxim of mechanism would thus remain regulative.<sup>52</sup> This suggests that, on Kant's account, the challenge raised by the apparently nonmechanistic character of living beings is a challenge that arises once the regulative status of the maxim of mechanism has already been accounted for. It is a challenge that persists even if we agree that the mechanistic maxim is a purely regulative guide for the study of nature.<sup>53</sup>

Second, the proposed reading of the maxim of mechanism also offers an answer to the worry that we may not be able to ascribe a determinate or objective status to mechanical explanations that are based, at least in part, on the employment of this regulative maxim. For although, as we have seen, the proposed account agrees with McLaughlin in arguing that the maxim of mechanism is a regulative principle of reflective judgment, it disagrees with his reading regarding the grounds of the maxim's nonconstitutive status. In order to see how this difference between the proposed reading and McLaughlin's interpretation accounts for the fact that the former avoids the difficulties of the latter, we need to examine, in more detail, the role played by the mechanistic maxim in the formation of mechanical explanations. According to the reading proposed here, the maxim does not lead us to make use of a purely regulative idea in our reflections about nature. Instead, it guides us to employ a concept that is objectively valid within the realm of material nature. The maxim of mechanism thereby does not instruct us to make further reflective judgments. Rather, by asking us to think about nature as a whole as falling under mechanical laws, it leads us to determine our experiences of nature by means of the concept of mechanical causation. Despite the regulative status of the maxim, it thus enables us to formulate empirical explanations that have determining and not merely reflecting status.

In this respect, the mechanistic maxim contrasts with the second maxim introduced by Kant in the "Dialectic of Teleological Judgment" as a guiding principle in our reflections on nature: 'Some products of material nature cannot be judged as possible according to merely mechanical laws (judging them requires an entirely different law of causality, namely that of final causes).'<sup>54</sup> This maxim presents a second, conflicting, interpretation of the regulative principle of the unity of nature. It asks us to reflect about nature



as a whole as ordered not only according to mechanical but also according to teleological principles. As we have seen in Section 1, however, the idea of the purposiveness of nature is a purely regulative idea. Kant therefore holds that we cannot ascribe purposiveness to nature itself, but can merely regard nature *as if* it were purposive. Teleological considerations can thus offer regulative reflections about nature without, however, enabling us to make determinate statements about the purposiveness of nature itself. By contrast with objective claims about the mechanical causes of nature, teleological judgments will thus always remain analogical considerations without any claim to objective validity. As analogical judgments, these reflections may then work indirectly as heuristic means of finding out about further mechanical causes. They may lead us to discover mechanical relations between parts of matter that were originally regarded as purposively related. The teleological judgments themselves, however, will always remain reflective. This contrast highlights, I suggest, that although both maxims present regulative rather than constitutive principles for the consideration of nature, only the mechanistic but not the teleological maxim directly instructs us to employ a concept that has objective validity. Application of the concept of mechanical causation, but not that of purposiveness, can thus lead to genuine explanations of nature.<sup>55</sup>

## V. CONCLUSION: CAUSAL KNOWLEDGE AS APPLIED RATIONAL COGNITION

In this chapter, I have thus argued that there is a crucial continuity between Kant's principle of causality in the *Critique of Pure Reason* and his maxim of mechanism in the *Critique of Judgment*. Rather than standing in conflict with one another, the constitutive principle of causality and the regulative maxim of mechanism provide necessary conditions for knowledge of the causes that determine change in material nature. Focusing on Kant's regulative conception of mechanical causation in the third *Critique* thus presents us with an account of causal knowledge that is essentially dependent on both constitutive and regulative principles.

In the *Metaphysical Foundations*, Kant characterises the resulting knowledge of mechanical causes as an 'applied rational cognition' and distinguishes it from the pure and apodictically certain part of knowledge, or 'science proper'.<sup>56</sup> This difference is grounded in the fact that empirical but not a priori statements comprise judgments about that which is independent of us and the particular character of which is therefore left undetermined by the a priori concepts of the understanding. Since judgments about that which is left undetermined by our a priori concepts must nevertheless be conceptual, empirical knowledge presupposes the availability of empirical concepts and principles that characterise the particular, contingent character of the given. And since, furthermore, these empirical



concepts cannot be read off from given sense perceptions—since the given is not yet conceptual but must be conceptualised by us—we must provide empirical concepts and principles ourselves.

As we have seen, we can do so only by reflecting about nature, guided by the assumption that nature can indeed be ordered by empirical concepts and laws. By constructing conceptual schemes, by means of which we make sense of that which is given to us in experience, we thus read systematicity and lawfulness into nature. The idea of a fully systematic knowledge of natural causes, however, is only something we can strive for, yet can never fully achieve. Further experience may always lead us to revise our conceptual schemes. ‘[I]n natural science’, as Kant therefore argues, ‘there is endless conjecture, and certainty is not to be counted upon’.<sup>57</sup> Despite such lack of certainty, however, we have reason to formulate and investigate the conjectures of science. Although, according to Kant, a regulative principle is necessary for the formation of causal explanations, we are therefore at least justified in *aiming* at causal knowledge of the world.<sup>58</sup>

## NOTES

1. Cf. *Critique of Pure Reason* (CPR), A189ff./B232ff. References to Kant’s texts are made by citing the volume and page number of *Kants gesammelte Schriften*, Preussische Akademie der Wissenschaften (ed.) (Berlin: Walter de Gruyter, 1902ff.), with the exception of the CPR, which is referred to by citing the page numbers of the original A and B editions. If not otherwise indicated, translations are from *Critique of Pure Reason*, trans. Norman Kemp Smith (London: Palgrave, 1929, first published 1781/1787); *Critique of Judgment*, trans. Paul Guyer and Eric Matthews (Cambridge: Cambridge University Press, 2000, first published 1790); and Immanuel Kant, *Metaphysical Foundations of Natural Science*, trans. Michael Friedman (Cambridge: Cambridge University Press, 2004, first published 1786).
2. This is not surprising, for the precise argument of Kant’s proof for the causal law is far from obvious and interpretative issues abound. A question that has dominated discussions of the ‘Second Analogy’ is the debate between a weak and a strong interpretation. Cf. note 41.
3. *Critique of Judgment* (CJ), V 360.
4. Reinhard Löw argues for this interpretation in *Philosophie des Lebendigen: Der Begriff des Organischen bei Kant, sein Grund und seine Aktualität* (Frankfurt am Main: Suhrkamp, 1980), 204ff.
5. Cf. Peter McLaughlin’s discussion of Löw in his *Kant’s Critique of Teleology in Biological Explanation* (Lewiston, NY: Edwin Mellen Press, 1990), 143f.
6. Cf. the discussion in Section 3.
7. CPR, B232.
8. CPR, B234.
9. CPR, B232.
10. CPR, A206/B251. Already in the *Dreams of a Spirit Seer* Kant calls any attempt to explain nature by means of immaterial principles a ‘resort of lazy reason’ (II 331). It is therefore a mistake to claim that the transcendental principle of causality leaves open the possibility of a different kind of causality. Two authors who argue for this possibility are Bernhard Rang,



“Naturnotwendigkeit und Freiheit. Zu Kants Theorie der Kausalität als Antwort auf Hume”, *Kant-Studien*, 81 (1990): 24–56; and Georg Sans, *Ist Kants Ontologie naturalistisch? Die “Analogien der Erfahrung” in der “Kritik der reinen Vernunft”* (Stuttgart: Kohlhammer, 2000).

11. CPR, A533/B561, translation amended.
12. CJ, V 386.
13. CJ, V 360.
14. CJ, V 386.
15. Cf. CJ, V 385ff.
16. CJ, V 386.
17. Ibid.
18. I discuss this conflict in Angela Breitenbach, “Two Views on Nature: A Solution to Kant’s Antinomy of Mechanism and Teleology”, *British Journal for the History of Philosophy*, 16 (2008): 351–69; and in *Die Analogie von Vernunft und Natur: Eine Umweltphilosophie nach Kant* (Berlin/New York: Walter de Gruyter, 2009), Chs. 5 and 7. I shall come back to a consideration of the status of these two principles in Section 4.
19. McLaughlin, *Kant’s Critique of Teleology*, 141ff. McLaughlin refers to A. C. Ewing, *A Short Commentary on “Kant’s Critique of Pure Reason”* (Chicago: University of Chicago Press, 1938, especially pp. 227f.) as an early pronouncement of this view.
20. Ibid., p. 153.
21. Ibid., p. 152.
22. Ibid.
23. Ibid., p. 166.
24. In a later paper, McLaughlin acknowledges that the necessity of Kant’s mechanistic interpretation of causality remains obscure (“Newtonian Biology and Kant’s Mechanistic Concept of Causality”, *Kant’s Critique of the Power of Judgment*, edited by P. Guyer, 209–218, 215f.). He argues that Kant’s characterisation of the human understanding could be explained by the fact that it relied on the actual methodology of the mechanistic science of his time. It remains unclear, however, how the empirical methods of science could justify what is characterised by Kant as a subjective condition of the human mind.
25. CJ, V 360.
26. CJ, V 410.
27. John H. Zammito, *The Genesis of Kant’s Critique of Judgment* (Chicago: University of Chicago Press, 1992), 223f., points out a similar difficulty as an objection to the interpretation of the concept of mechanism proposed by John D. McFarland, *Kant’s Concept of Teleology* (Edinburgh: Edinburgh University Press, 1970). Similar to McLaughlin, McFarland argues that the concept of a mechanical explanation presents a merely subjective idea: the ‘quantitative abstraction’ that mechanical explanations rely on ‘are not found in nature, but are introduced into nature for our own convenience in explaining it’ (p. 31).
28. *Metaphysical Foundations* (MFNS), IV 543. Cf. the discussion of Kant’s second of law of mechanics in Michael Friedman, “Matter and Motion in the Metaphysical Foundations and the First Critique: The Empirical Concept of Matter and the Categories”, in *Kant and the Sciences*, edited by E. Watkins (Oxford: Oxford University Press, 2001). 53–69; and Konstantin Pollok, *Kants “Metaphysische Anfangsgründe der Naturwissenschaft”: Ein Kritischer Kommentar* (Hamburg: Felix Meiner, 2001), 414ff.
29. MFNS, IV 536.
30. Hannah Ginsborg, “Two Kinds of Mechanical Inexplicability in Kant and Aristotle”, *Journal of the History of Philosophy*, 42 (2004): 33–65, 43.



31. CJ, V 408.
32. CJ, V 407. I here agree with Ginsborg, 'Two Kinds of Mechanical Inexplicability', that the mechanism of nature should be read in terms of the general laws of mechanics that Kant sets out in the *MFNS*. I do not agree with Ginsborg's further claim, however, that the mechanism of nature has nothing to do with causality. Rather, as I argue here, it is the necessity of a mechanical interpretation of causality in the context of material nature that accounts for the necessity of mechanical explanations in science. I have argued for this interpretation in more detail in Angela Breitenbach, "Mechanical Explanation of Nature and Its Limits in Kant's Critique of Judgment", *Studies in History and Philosophy of Biological and Biomedical Sciences*, 37 (2006): 694–711.
33. Watkins suggests as much in "The Antinomy of Teleological Judgment", in *Teleology. Kant Yearbook* 1, edited by D. Heideman (2009), 197–221, 205.
34. CJ, V 386.
35. Ibid.
36. Ibid.
37. CPR, A642ff./B670ff.
38. CPR, A653f./B681f., italics mine.
39. CJ, V 185. Kant spells out this worry further both in the first, unpublished, introduction (especially FI, XX 203ff.) and in the second, published, introduction to the *Critique of Judgment* (especially V 179ff.).
40. Different commentators have pointed out that Kant mentions different aspects of nature that are left open by the transcendental principles of human understanding. Thus, Kant refers to the empirical conceptualisability of nature, the determination of nature by empirical laws, the systematicity of such empirical laws as well as their necessity. Cf. Henry E. Allison, *Kant's Theory of Taste: A Reading of the Critique of Aesthetic Judgment* (Cambridge: Cambridge University Press, 2001), 30f.; Paul Guyer, *Kant's System of Nature and Freedom: Selected Essays* (Oxford: Clarendon Press, 2005), 48f.; and Rachel Zuckert, *Kant on Beauty and Biology: An Interpretation of the Critique of Judgment* (Cambridge: Cambridge University Press, 2007), 27ff. I agree with Allison that we can find some coherence in Kant's various formulations if we keep in mind that it is 'the essential function of reflective judgment . . . to find universals for given particulars' and thereby to unify particular experiences into a systematic whole (ibid.).
41. I believe that this problem arises for both the weak and the strong interpretation of the "Second Analogy". According to the weak reading, Kant's argument establishes that all of nature has some cause but not that nature is determined by causal laws. Cf., e.g., Lewis White Beck, *Essays on Kant and Hume* (New Haven/London: Yale University Press, 1978), 111–29; Gerd Buchdahl, *Metaphysics and the Philosophy of Science: The Classical Origins: Descartes to Kant* (Oxford: Basil Blackwell, 1969); and Henry E. Allison, *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy* (Cambridge: Cambridge University Press, 1996), 80–91. According to the strong reading, by contrast, the "Second Analogy" proves not only that every event has a cause but also that the same types of cause have the same types of effect. Cf., e.g., Arthur Melnick, "The Second Analogy", in *A Companion to Kant*, edited by G. Bird (Oxford: Blackwell, 2006), 169–81; Paul Guyer, *Kant and the Claims of Knowledge* (Cambridge: Cambridge University Press, 1987); Michael Friedman, "Causal Laws and the Foundations of Natural Science", in *The Cambridge Companion to Kant*, edited by P. Guyer (Cambridge: Cambridge University Press, 1992), 161–99. The problem that remains open for both interpretations, I believe, is that of *discovering* the particular causal laws.



42. CPR, A647/B675.
43. The relationship between the role of reason in the CPR and that of reflective judgment in the CJ has been the object of much debate. Cf. for instance the discussion in Joachim Peter, *Das transzendente Prinzip der Urteilskraft: Eine Untersuchung zur Funktion und Struktur der reflektierenden Urteilskraft bei Kant* (Berlin/New York: Walter de Gruyter, 1992), part 2; and Suma Rajiva, "Is Hypothetical Reason a Precursor to Reflective Judgment?", *Kant-Studien*, 97 (2006): 114--26. For the purpose of this chapter, I shall assume that in the CJ we can still regard the idea of unity as an idea of reason, an idea, however, that is projected onto nature by our reflective activity of judgment. It is particularly controversial whether Kant regards the principle of the systematic unity of nature as a transcendental principle only in the CJ or also in the CPR. According to Rolf-Peter Horstmann, *Bausteine kritischer Philosophie: Arbeiten zu Kant* (Bodenheim bei Mainz: Philo, 1997), Chs. 5–6, and Guyer, *Kant's System*, Chs. 2–3, Kant's position changes between the first and third *Critiques*. By contrast, Allison, *Kant's Theory of Taste*, and Ido Geiger, "Is the Assumption of a Systematic Whole of Empirical Concepts a Necessary Condition of Knowledge?", *Kant-Studien*, 94 (2003): 273–98, argue convincingly that the CJ develops what was already implied in the CPR.
44. Cf. CJ, V 179.
45. CJ, V 184f.
46. FI, XX 209.
47. CJ, V 386.
48. Ibid.
49. MFNS, IV 543.
50. CJ, V 387.
51. CJ, V 386.
52. According to Marcel Quarfood, 'the plurality of special laws harmonising in sustaining the organism presents a complexity which at least with respect to the "peculiarity" of the human vantage point is infinitely too vast for the mechanical mode of explanation to be anything more than just a regulative principle' (*Transcendental Idealism and the Organism: Essays on Kant* [Stockholm: Almquist & Wiksell, 2004], 205). According to the reading I have proposed here, however, it is not the particular complexity of organisms that makes the 'mechanical mode of explanation' purely regulative. Rather, it is the claim that all of nature can be so explained that must be regarded as a purely regulative guide to our mechanical explanation of particular natural objects.
53. This is why the conflict between a mechanistic and a teleological maxim remains, according to Kant, even if these maxims are understood as purely regulative principles.
54. CJ, V 387.
55. This reading does not yet provide a solution to the conflict raised by these principles. It does, however, suggest that the resolution of the antinomy of judgment must take into account the different status of judgments made by employing the concept of mechanical causation, on the one hand, and that of purposiveness on the other. For a suggestion of how to read this resolution, see Breitenbach, "Two Views on Nature".
56. MFNS, IV 468.
57. CPR, B433.
58. I would like to thank Nick Jardine, Marina Frasca-Spada, Sasha Mudd, and the participants at the BSHP Conference 2008 for helpful comments on earlier versions of this chapter.



# 11 Regularities All the Way Down

## Thomas Brown's Philosophy of Causation

*Stathis Psillos*

### I. INTRODUCTION

Thomas Brown (1778–1820) was one of the tail-enders of the Scottish Enlightenment. He shared with Dugald Stewart (1753–1828) the chair of moral philosophy at the University of Edinburgh from 1810 until his premature death in 1820. He is sometimes classed with the Scottish commonsense philosophers and, to some extent at least, his basic philosophical principles were akin to those of the commonsense philosophy. He did, for instance, forfeit the issue of the justification of some of our most basic beliefs and rested them, instead, on their being intuitively irresistible; in particular, he thought that some of our most basic beliefs could be seen as permanent principles of human nature—a claim made popular by Thomas Reid. Based on his theory of the workings of the human mind—which was developed in a course of lectures on the philosophy of mind presented at the University of Edinburgh and appeared posthumously as a book titled *Lectures on the Philosophy of the Human Mind*—some philosophers and psychologists have characterised him as an ‘associationist’.<sup>1</sup>

Brown's main contribution to the philosophy of causation was his book *Inquiry into the Relation of Cause and Effect*, published in 1818.<sup>2</sup> This is, actually, the third (substantially enlarged and developed) edition of his little book titled *Observations on the Nature and Tendency of the Doctrine of Mr. Hume Concerning the Relation of Cause and Effect*, which was published in Edinburgh in 1805 and made a second edition in 1806.<sup>3</sup> This little pamphlet was motivated by the so-called Leslie affair. When the chair of mathematics became vacant in the University of Edinburgh, John Leslie applied for it. But the city ministers were vehemently opposed to this appointment, arguing that Leslie was a defender of Hume's view of causation, which was taken to deny the existence of an almighty God. Brown's pamphlet aimed, among other things, to save a Humean view of causation from the charge of atheism.

This work on causation was Brown's defining philosophical moment. His views were discussed in the nineteenth century, but more recent philosophers and historians of causation have paid very little attention to



them. The only relatively recent paper that aimed to discuss Brown's views is John Mills's.<sup>4</sup> The truth is that Brown's views constitute a carefully crafted attempt to develop a theory of causation that differed substantially from those offered by the two main figures of his era: David Hume and Thomas Reid.

Brown took it that causation as it is in nature is regular and invariable succession but he argued against Hume's theory of the locus of the idea of necessary connection. However, Brown aimed to save Hume from the then popular criticism that he denied that we have an idea of power. At the same time, he wanted to show that Hume was wrong in allowing that talk about powers might be taken to be talk about something different from the very invariable succession of events. In terms of the current New Hume debate, Brown did perceive that Hume might be taken to be a sceptical causal realist of a sort.<sup>5</sup> Though Brown wrote after Hume and was deeply influenced by him, we might say that he was the first defender of a pure and simple regularity theory of causation. If the friends of a new Hume are right about Hume, the doctrine that has been called Humeanism had better be renamed *Brownianism*.<sup>6</sup>

The immediate intellectual milieu within which Brown developed his theory of causation was dominated by Reid's power-based account of causation (cf. 1863, Essay 1). Brown's philosophy was a revolt against powers. We shall characterise his view as an identity theory of powers. Brown did not deny that we can meaningfully talk about powers. He did deny, however, that there is anything like a causal nexus or a tie between distinct existences in virtue of which they fall under patterns.

It is remarkable that Brown (1822, 212) ends his book on causation with the claim that Reid and Hume share belief in power as, ultimately, something distinct from invariable sequences of events in nature. They differ, among other things, in how this belief arises. On Brown's view (1822, 203), however:

*Power* is only a shorter synonymous expression of *invariableness of antecedence*: and the invariableness is not a thing separable or distinguishable from the antecedents and consequents themselves. In all the changes which the substances in nature undergo, the substances themselves alone have real existence; and what we term Power, in the anticipation of any future change, is itself the antecedent substance, or it is nothing.

This chapter offers a systematic account of Brown's philosophy of causation in relation to Reid's and Hume's. It will be argued that Brown had a very sophisticated theory of causation that aimed to answer all three of the following questions: (i) What is causation as it is in the world? (ii) What do we mean when we talk about causation? (iii) How does belief in causation arise? In the course of answering these questions, Brown vehemently denied



that there is need to posit anything metaphysically more robust than uniformity to account for all there is to causation. At the same time, he developed a challenging theory of the grounds for belief in uniformity. The bottom line of this is that if we allow for no Archimedean points in epistemology (that is, for beliefs we are entitled to without being able to justify them, according to some strict conception of justification), there is no way to avoid scepticism. Finally, in trying to show how a regularity view of causation does not fall prey to some objections levelled against it, Brown gave to this view a sophisticated twist, arguing that though causation is a species of regularity, what regularities underpin certain causal relations are not to be read off directly from the rough-and-ready descriptions of the causal relata.

## II. REID VS. BROWN

### II.1 Reid on Causation and Active Powers

Reid spoke freely of active powers and took it that (a) the very concept of power is simple and undefinable; (b) power is *not* something we either perceive via the senses or we are aware of in our consciousness (we are conscious only of the *operation* of power and not of the power itself); (c) power is something whose existence we infer by means of reason based on its operation; (d) power is distinct from its manifestation/exertion in that there may be unexerted powers; (e) the idea we have of power is relative, namely, as the conception of something that produces or brings about certain effects—hence, ‘our conception of power is relative to its exertion or its effects . . . *as something that has a certain relation to [an] effect*’ (1863, Essay 1, Chapter 1, 514); (f) power always requires a subject to which it belongs: it is always the power *of* something; the power that something *has*; (g) causation is the production of change by the exercise of power. Reid insisted that though we are not conscious of powers, we are conscious of their exertion when our own mental active powers are exercised, as when we decide to raise our hands.

Despite this rich conception of active powers, Reid did not think that, strictly speaking, there is causation *in* nature. He did think that the cause of a change is that which produces the change by the exercise of its powers. Since the very idea of exercising a power requires agency and there is no agency in nature, strictly speaking, there is no causation in nature. Indeed, Reid (1863, Essay 1, Chapter 6) insisted that properly understood, active powers require subjects that have intelligence and will to exercise them. Inanimate matter then can be no such subject. Only God—who is an ‘off-stage agent’—can be the cause possibly by means of secondary causes.

In an undated letter to James Gregory, Reid distinguished between the ‘strict and proper’ sense of ‘cause’ and the ‘lax and popular’ sense of



it. According to the first, causes are active powers to produce an effect.<sup>7</sup> According to the second sense, ‘a cause . . . means only something which, by the laws of nature, the effect always follows’. This second sense is akin to the view that causation is regular succession, though Reid took it that laws of nature are principles of necessitation. Even thus understood, the second sense of ‘cause’ is not enough for causation—which, properly understood, has to be *efficient* causation. He added: ‘I think natural philosophers, when they pretend to shew the causes of natural phenomena, always use the word in this last sense; and the vulgar in common discourse very often do the same’. In (1863, Essay 1, Chapter 6, 527), Reid explained that this subsumption under laws of nature does not constitute causation; nor does it amount to causal explanation. For him,

the laws of nature are the rules according to which the effects are produced; but there must be a cause which operates according to these rules. The rules of navigation never steered a ship. The rules of architecture never built a house.

Hence, a cause is something that has the power to bring about an effect *in accordance with the law*; but knowing the laws does not amount to knowing the causes. From all this, he drew the rather pessimistic conclusion that in spite of the fact that scientists have discovered a number of laws of nature, ‘they have never discovered the efficient cause of any one phenomenon’ (ibid.). Which, for Reid, is just as well since those scientists who understand what science is about and what the laws of nature are do not claim that science discovers (or aims to discover) causes. For Reid, as we have already noted, causation is tied to agency and laws of nature are not agents. As he (1863, Essay 4, Chapter 3) put it:

[laws of nature] are not endowed with active power, and therefore cannot be causes in the proper sense. They are only the rules according to which the unknown cause acts.

Reid was a vocal critic of the view that causation amounts to regular succession—a doctrine he associated with Hume. One of his chief points was that the regularity of succession could never lead us to the notion of cause unless we were already convinced that every event has a cause. This, for him, is a principle of the constitution of the human mind—a *first principle*—which is universal and basic and yet not a truth of reason. More concretely, he criticised Hume for having reduced consequence to mere sequence. The claim that was to become famous was that Hume’s doctrine—the regularity view of causation—implies the absurdity that the day is the cause of night and the night is the cause of day because they have constantly followed each other. As Reid (1863, 606) characteristically put it:



Nor is that always the cause of a phenomenon which is prior to it, and constantly conjoined with it; otherwise night would be the cause of day, and day the cause of the following night.

Though Reid based his account of causation on active powers, and though he insisted on a sharp metaphysical separation between the power and its manifestation, he admitted that powers cannot be observed and that it is only the regular sequence of events that can. As he (1863, Essay 4, Chapter 6, 617) put it: ‘We perceive one event to follow another, but we perceive not the chain that binds them together’.

## II.2 Brown on Regularity

It is precisely claims such as this that gave Brown the basis for his critique of a power-based account of causation. The sought-after chain of causation is a chimera, based—at best—on a metaphorical use of language. Insofar as there are loops in the ‘chain’ that links cause and effect, they are just intermediate steps in the regular association between events like the cause and events like the effect. Interestingly, for Brown the only sense of causation is Reid’s ‘lax and popular’ one. Laws of nature, he thought, are ‘the accustomed order of the sequences of the phenomena of Nature’ (1822, 56) and to ascribe a power to a thing is to place it in relation to a law of nature, namely, to an order of succession.

The clearest summary of Brown’s account of causation is this (1822, 21):

A cause, in the fullest definition which it philosophically admits, may be said to be, that which immediately precedes any change, and which, existing at any time in similar circumstances, has always been, and will be always, immediately followed by a similar change. Priority in the sequence observed, and invariableness of antecedence in the past and the future sequences supposed, are the elements, and the only elements, combined in the notion of a cause. . . . [P]ower . . . is only another word for expressing abstractly and briefly the antecedence itself, and the invariableness of the relation.

Brown insists that when we try to understand causation, we try to understand at least two things: first, what causation is; and second, how belief in it arises. Given, however, the popularity of power-based accounts of causation, Brown adds a further task, namely, to understand and unravel the sources of the ‘illusion’ that have led philosophers to think that there is something behind or beyond the regularity that enforces it.

Why, one might ask, is the thought appealing that causation involves regularity? Brown starts with some standard observations. One is that though there is a lot of change in the world (‘The world is a mighty system



of changes', 1822, 17), changes fall under regular patterns. When there are deviations from regular patterns, the natural tendency is not to deny regularity but rather to attribute them to interfering circumstances. When a regularity (e.g., All As are B) is denied, another (more complicated one) is affirmed (e.g., All As and Cs are D), since when present the interfering circumstances will give rise to a new effect. Brown claimed that had there not been regular patterns in the world, we would not have the concepts of causation or power. This is, at least partly, because the concept of causation arises in connection with activities such as action, planning, predicting, and controlling. The presence of regularity renders effective strategies (such as planning and prediction) possible. As Brown put it, it is because 'the future, when it arrives, we find to be only the past under another form' that we can materialise our wishes, fulfil our plans, and succeed in our actions.

Reid, as we have seen, would definitely disagree with tying the concept of causation with regularity, since he took it that we have at least a relative conception of power stemming from the exercise of our own will. But Brown, taking Hume's side, denied that there is anything like mental power distinct from uniformity; nor is there, according to him, any direct conception of mental power. Here is a nice summary of his views:

The theory of Power, then, seems to receive no additional light from consideration of mental energy, as exhibited in the bodily movements that depend upon the will; for we find, as before, only a sequence of two phenomena, that are believed to be, in the same circumstances, uniformly antecedent and consequent (1822, 40).

Regularity might be present in the world. It might be important for effective strategies. It might even be a sign of causation. Still, causation might be more than regularity. Brown is certainly aware of this problem. He wants to defend the strong view that regularity (invariable sequence) is constitutive of causation as it is in nature. Hence, he needs to block arguments to the effect that causation might well have some other essential characteristic in virtue of which it is exemplified in regular sequences of events and hence that invariable sequence is *merely* a sign of causation—which is actually something else. Brown's strategy was precisely to show that regularity is all there is to causation; it 'is itself the only essential circumstance of causation' (1822, viii). To achieve this, he advanced two kinds of argument. On the one hand, he developed a positive argument against powers—advancing what might be called the *identity-theory of powers*: powers are nothing but the regularity, the uniformity of sequence. On the other hand, he articulated a number of negative arguments aiming to show 'the sources of various illusions' which have led philosophers to posit powers and to consider causation something more 'mysterious' than regularity.



### III. POWERS UNMASKED

Powers, according to Brown and the then (and now) standard conception, were supposed to be inherent in objects and yet distinct from them; they were supposed to account for the efficiency of causation. But Brown forcefully denied that between the cause and the effect there is something else (an ‘intermediate tie’ or an ‘invisible bondage’) that connects them or binds them together; in particular something of a radically distinct metaphysical nature.

#### III.1 Brown on Properties

Being a nominalist, Brown has reasons to suspect this hypostatisation of powers, anyway.<sup>8</sup> His general view (cf. 1822, 24) was that there were only individual substances—that is, particulars. He denied that there were universals. General terms, like *Man* or *Animal*, were classificatory schemes introduced for convenience. They do not refer to anything other than classes of resembling particulars. They do not denote any separate substances (like substantial universals). Similarly, predicates do not denote nonsubstantial universals that exist, somehow, independently of the individual substances and are, somehow, possessed by (instantiated in) them. It is, as he put it, a ‘monstrous species of realism’ that has led a number of philosophers to hypostatise universals. Perhaps the worst case of this monstrous realism is the hypostatisation of powers. It has fostered the thought that there are necessary connections in nature, where there are none.

So, according to Brown, there are no powers distinct from substances. In any case, to attribute powers to substances is to attribute properties or qualities to them. To say, for instance, that water has the power to melt salt is to say that water has the *property* to melt salt; or that it is a *quality* of water to melt salt. For Brown, the properties of substances, for example, the greenness of the emerald, or the yellowness of gold, or the specific gravity of gold, are understood as conditional attributions to substances. To say that a substance *x* has the property *P* is to say that if *x* is in circumstances *C*, then effect *E* follows. The effect might be a successive state of the very same substance (as, for instance, in inertial motion) or a change in a different substance (as, for instance, in the heating of a body). So to attribute properties to a substance is to place it in a relation of cause and effect, either to other substances or to later states of the same substance. Powers are treated in a similar fashion. Talk of powers, Brown argues, merely signifies what a substance does under certain circumstances—for example, that when water is poured on salt, what was previously a crystalline substance gets liquefied. More generally, to say that a substance has certain powers (that is, properties) is ‘to consider it as existing in a variety of circumstances, and to consider at the same time all the changes that are or may be in these circumstances its immediate effect’ (1822, 20). Here is a nice example that summarises Brown’s views:



In the beautiful experiment of the prismatic decomposition of light, for example, the refracting power of the prism is not any thing separate or separable from it, more than its weight or transparency. There are not a prism and transparency, but there is a prism giving passing to light. In like manner, there are not a prism and refracting power, and coloured rays, but there are a prism and rays of various colours which we have perceived to be deflected variously from their original line of direction, when they approach and quit the lens, and which we believe will, in the same circumstances, continually exhibit the same tendency (1851, 39)

Brown is unclear as to whether this view of properties extends to the primary qualities of a substance, but the overall tone of his argumentation suggests that it does. He favours a worldview according to which there are (simple) substances and they stand in certain spatiotemporal relations to one another and fall under certain patterns of invariable succession: whenever this-type of thing happens, that-type of thing follows. The laws of nature are these general patterns, namely, the regularities. The properties attributed to the substances simply codify the ways that the substances are related to each other. So properties (and powers) are conceived of as a relational net, the nodes of which are substances. To ascribe powers to a substance is to consider it in various circumstances and to then consider what changes, as a matter of fact, follow. The powers thereby attributed to a substance are its properties and its qualities. They are not distinct from the substance ('superadded to it') but instead it is a way to view the substance itself 'in relation to various changes that take place when it exists in peculiar circumstances'. This is not an eliminativist view of properties. Brown accepts that a substance (matter, as he would put it) without qualities 'seems to be a contradiction in terms' (1822, 70). But for him the very conception of a substance (matter) with qualities requires taking for granted that this substance is the invariable antecedent of certain changes. This is an important point. For Brown, to attribute qualities to a substance is precisely to place this substance in a network of regularities (or laws, if you like) that relate this substance with others as cause and effect:

All this regularity of succession . . . is assumed in our very notion of substances, as existing (1822, 70).

In his *Lectures*, Brown summarised his metaphysics in a very interesting way. When we ask of an object (a substance) *what is it?*, the answer is to place it in space and time and to consider (a) the objects that coexist with it in space (that is, its *constitution*) and (b) the objects that are related to it in time, that is, its causes and its effect: 'all the series of changes, of which it forms an invariable part, the objects to which it is related as antecedent or consequent' (1851, 36). Let us picture this as a vast spatiotemporal mosaic (to use a well-known Lewisian expression) of objects and let us include in



this mosaic all the regularities (regular patterns of succession) of which these objects partake. This vast mosaic determines all there is in the world. In particular, it determines all causal facts; and it fixes all powers that there are in the world.

Criticising Brown, James Peterson invited us to consider a possible world in which ‘there should never be two causes alike and therefore never two events alike’.<sup>9</sup> In that world, he claimed, there would still be causation: ‘every event in that world would have its cause as surely as in this world’. This is exactly what Brown denies. This would be a world of *casual* and not of causal sequences.

### III.2 An Identity Theory of Powers

Brown (and the Reverend David Welsh, his biographer and follower) warn us that moving from the claim that powers are nothing but X, to the claim that powers are *nothing*, is a fallacy. To say, in particular, that power is nothing more than ‘invariableness of antecedence’ is not to say it is nothing. Welsh draws a nice parallel between the powers of a substance and a net.<sup>10</sup> A net is so constructed that it retains objects of certain sizes and allows the passing of other objects. These are powers of the net. But they are wholly constituted by the structure of the net and the relations it has to other objects (e.g., the fish). Different powers attributed to a substance then are simply nothing other than different relations to which this substance stands to different objects. Even then, however, Brown warns us not to proliferate powers. Heat does *not* have the power to produce a certain sensation of warmth and a distinct power to melt wax. There is simply the heat in relation to two distinct objects, my body and the wax.

Arguably, Brown advanced an identity theory of powers, according to which, ‘power is [the] uniform relation [between cause and effect] and nothing more’ (1822, 26). Hence to ascribe a power to an object is nothing, but to assert that in similar circumstances, it will do similar things. Apart from the general philosophical motivation noted earlier, this theory is based on a number of arguments, aiming to show that there is no need to posit powers over and above the regularities.

*First*, powers are mere abstractions (cf. 1822, 19ff.). A causal sequence is a concrete sequence between events. It is *causal* in virtue of the fact that this sequence is invariable (it exhibits regularity of order), namely, its antecedent (the cause) has been followed, is followed and will be followed by its consequent (the effect). When we consider this relation (*this* is always followed by *that*) abstractly, that is independently of the particular circumstances in which it takes place, we render the ‘—is always followed by—’ as ‘—has the power to—’. This move is supposed to unravel the form of causation, namely, what several concrete causal sequences have in common and in virtue of which they are causal. This move, for Brown, is akin to the hypostatisation of substantial forms and suffers from exactly the same



problem: it converts an abstraction to reality, thereby creating the further problem to explain what this kind of new entity is and does. Power, then, is merely the very invariableness of the order of succession, abstractly understood (See also 1851, 35).

*Second*, powers are the products of double vision (cf. 1822, 28–29). There are substances and they stand in causal relations to each other (that is, in relations of invariable succession). If we knew all these invariable sequences, we would know everything there is to know about what causes what. If we then, based on this kind of knowledge, we added that these substances have the *power* to produce certain changes, we would not gain any further information about the world. If we thought of power as distinct from these invariable sequences, that is, as something over and above the invariable sequences, something that an object possesses and in virtue of which causes whatever it does, then it would be possible that we could have information about invariable sequences without having knowledge of a single power.

*Third* (and relatedly), powers do not explain the regularities. The existence of regularities in nature is not rendered ‘less wonderful’ by an appeal to powers. (cf. 1851, 36). Actually, since powers can exist unexerted, there may be no regularities.

*Fourth*, positing powers as distinct existences arises out of the confusion between casual and causal sequence (cf. 1822, 29). A single spatiotemporal sequence of events is not enough for causation (this is a *casual* sequence). It is only ‘similarity of sequence’ that underpins causation. Powers may be thought as necessary to bridge the gap between a casual sequence and a causal one, but there is no such need provided that causal sequences are understood as invariable sequences: ‘to know events as invariably antecedents and consequents is to know them as causes and effects; and to know all the powers of every substance, therefore, would be only to know what changes or events would, in all possible circumstances, ensue, when preceded by other changes or events’ (1851, 40).

*Fifth*, powers are not needed for the explanation of action (cf. 1822, 56–57). Action amounts to making a difference. An object does not act on anything if its presence or absence makes no difference to anything. But this difference-making can be understood as invariable sequence. Objects that act and are acted upon (that is, causes and effects) are ‘truly, in certain circumstances, the reciprocal and immediate antecedents and consequents, in a series of changes’ (1822, 56–57).

### III.3 The Sources of Illusion

When it comes to the sources of illusion which have led philosophers to posit powers, Brown argued that they are of three kinds. The *first* (cf. 1822, Second Part, Section II) relates to language and the use of a number of metaphorical phrases when we think of causation, such as ‘connection’ or



‘bond’. For Brown, this metaphorical use of language has led us mistakenly to assume that there is something other than the ‘regularity of succession’ that constitutes causation. The *second* (cf. 1822, Second Part, Section III) relates to a folk metaphysical belief that there can be latent powers, that is, that things have the power to act in certain ways even if they are not acting—powers being that in virtue of which they can bring about the effect even if they do not. Here, Brown takes a hard line. He stresses that, strictly speaking, it makes no sense to talk about powers when they are not exerted, that is, to talk about powers as being latent in the intervals of their exertion. His argument for this claim can be reconstructed as follows. Powers are supposed to be the producers of change. When a power is exerted, it produces a change. When, however, a power is not exerted, there is no change whatsoever brought about by it. So an unexerted power produces no change and hence it cannot be a power. The *third* (cf. 1822, Second Part, Section IV) stems from a folk epistemological view that since we do not know how and why a cause produces the effect, there must be something intermediate and distinct from the cause, which is unknown to us, that is to say, a power of the cause to bring about the effect. Here, Brown insists that there is nothing of a distinct metaphysical kind to be found between the cause and the effect—just more of the same stuff: invariable antecedents to the effect.

It is noteworthy that Brown’s point is not that powers are suspicious because they are unobservable. He allows that there can be things we do not observe and that ‘we see only parts of the great sequences that are taking place in nature’ (1822, 92). His point is that even if we could see much more than we do, we could see more links between the cause and the effect, that is, more and more invariable antecedents, but we could not find the metaphysically distinct entity we call ‘power’, that ‘mysterious unintelligible something, between entity and nonentity, which we now conceive it to be, or rather, of which we vainly strive to form a conception’ (1822, 92–93).

### III.4 Efficiency in causation

One worry that one may have about a regularity theory of causation concerns the modal force that causation is supposed to have. Brown did not deny that causation involves efficiency. As he put it: ‘Causation is efficiency; and a cause which is not efficient, is truly no cause whatever’ (1822, 59). But he did not think that efficiency has anything to do with a productive power. Criticising Malebranche’s distinction between efficient causes and physical causes, he argued that (a) physical causes are efficient causes; and (b) physical causes are immediate and constant antecedents. Insofar as a causation is invariable succession, insofar, that is to say, that *C* is the cause of *E* is to say that *E* is an invariable consequent of *C*, prefixing the word ‘cause’ with the word ‘efficient’ or ‘physical’ is superfluous. The idea that efficiency is something distinct from invariable succession is fostered by the feeling that if *C* causes *E*, the cause will never appear without being



followed by the effect and the effect will never appear without being preceded by the cause. Brown is ready to allow for this claim and he suggests that it might be understood in terms of counterfactual conditionals. As Welsh notes, the notion of cause gives rise to counterfactuals of the form: 'if the cause had not existed, the effect would not have taken place'.<sup>11</sup> But Brown was careful not to think that causation could be defined in terms of counterfactuals. Indeed, Brown discussed in some details (in endnote A of his 1822) Hume's famous 'other words' that he appended to his first definition of causation, namely,

[W]e may define a cause to be *an object, followed by another, and where all objects, similar to the first, are followed by objects similar to the second*. Or, in other words, where, if the first object had not been, the second never had existed (1974, 76).

These are not 'other words', of course, and Brown picks on this. His point is twofold. *First*, this counterfactual account is admissible if there are simple trains of events, where there is no overdetermination; but *second*, the possibility of overdetermination cannot be excluded. Here is his example. Take a piece of iron being attracted by a magnet. The very same effect in terms of the motion of the piece of iron could have taken place if the piece of iron were held by a hand and was moved by it towards the magnet. Or take Welsh's example: a sword might enter a vital part of a body and cause death; but at the same time there can be another cause in operation.<sup>12</sup> The point of all this, of course, is that in light of the possibility of overdetermination, a counterfactual *definition* of causation such as that noted earlier would be inadequate. As Brown put it: 'the first object might not have been, and yet the second might have existed'. In the end, Brown takes it that it is enough to characterise causation as invariableness of sequence, since he takes it that it is not necessary for defining causation to take into account what it might or might not have been, in other circumstances where the antecedent was different from the actual one. For Brown, the locus of whatever modal force causal claims have is found in their future extendability, and in particular in our belief that the sequence is (and it will be) invariable.

#### IV. INTUITIVE BELIEFS

Brown agreed with Reid that not all beliefs are based on either reason or experience. Some beliefs are intuitive. In fact, some beliefs present themselves with such a force that they are irresistible. These are beliefs or opinions, which it is impossible for us not to hold, because of 'the very constitution of our nature' (1822, 149). Of this kind is, for Brown, the belief in regularity. It stems from a 'peculiar tendency of our constitution, which we must take for granted' (1851, 35). As we shall see in detail in the next section, Brown took to heart Hume's claims that causal beliefs (and belief in causation in general)



cannot be the product of reasoning based on past experience. Unlike Hume, however, Brown found little consolation in the observation of similarity of sequence, that is, in the observation of constant conjunction between event-types, as the source of this belief. There is, to be sure, constant conjunction and similarity of sequence in nature, but this cannot be the source of our belief in regularity because, as Hume himself noted, there is no qualitative difference between what is observed in a single sequence and what is observed in a thousand of them. The observation of single sequence of events is as embedded in the past as the observation of a thousand of those and the belief in regularity concerns the future as well as the past (and the present). This leaves only one option available, namely, that the source of this belief is intuition. When we believe that an event caused another, we believe that their relation is permanent, namely, that when similar circumstances arise, the same cause will be followed by the same effect. This belief does not rely on argument but it is impossible for us not to have it.

The distinctive marks of intuitive beliefs are: they are universal, immediate, and irresistible. The principles which are the content of intuitive beliefs are 'first truths', such as our own personal identity through time, or the reliability of memory, or the uniformity of nature. Belief in them is direct; noninferential. It is not a product of reasoning. In his *Lectures*, Brown went as far as to claim that principles such as the aforementioned are 'so irresistible in evidence as to preclude the possibility of denial'. That's clearly too strong a claim, since (as Hume himself noted) denying principles such as these is not self-contradictory. But Brown was influenced by Reid in accepting that unless some first principles are taken as self-evident starting points of inquiry, there is no possibility of inquiry and the road to scepticism is open. In highlighting this point, he takes these principles to be Archimedean points of inquiry. To deny them is 'to set [our] feet upon the air rather than on the ground . . . and to throw away the single fulcrum on which [our] lever rests and from which alone all its power is derived' (1851, 82). Where he actually disagreed with Reid was not the inevitability of such first principles (which Brown took them to be 'a necessary part of our intellectual constitution', 1851, 78) but their extent. Brown thought they should not be multiplied beyond necessity.<sup>13</sup>

Given his identity theory of power, Brown can easily claim that 'The belief in power is an original feeling, intuitive and immediate on the perception of change; not borrowed from any *resemblances* in the transitions of thought' (1822, 199).

## V. BROWN VS. HUME

Brown's account of Hume's views of causation is both a qualified defence and a critique. The defence has mostly to do with blowing away a widespread misreading of Hume that has followed the inception of the *Treatise*,



namely, that Hume takes the concept of power to be meaningless. Brown agreed with Hume on the following three principles. *First*, causation is not a relation that can be known a priori; *second*, reason cannot lead to the establishment of causal relations even when aided by experience; *third*, therefore, causation is only an object of belief. But he disagreed with two more principles of Hume's. *Fourth*, belief in causation arises only after observations of constant conjunctions; and *fifth*, this belief is marked by a transition of the mind from the idea of the cause to (an even more vivid) idea of the effect.

## V.1 The Third Factor

Brown (1822, Part IV, Section VI) starts with a masterly discussion of one of Hume's central arguments in the *Treatise*. According to a common reading of Hume, his argument is this:

- (A)
- We have no idea which is not copied from impressions.
- We have no impression of power.
- Therefore, we have no idea of power.

The major premise of this argument is undeniably Hume's own—it is his major methodological maxim, which ties the presence and meaningfulness of ideas to impressions. So: no idea can be contentful, unless it corresponds to a (prior) impression. The minor premise of (A) has been attributed to Hume, not altogether without justification—since he does deny that there is anything corresponding to power that can be *observed* in a causal sequence of events. But Brown takes it that (A) is not Hume's argument—so Hume is not committed to its conclusion. Hume's argument, instead, is:

- (B)
- We have no idea which is not copied from impressions.
- We have an idea of power.
- Therefore, there must be an impression from which the idea of power is derived.

Indeed, (B) is much closer to what Hume had been doing in the *Treatise*. After all, Hume never denied that part of the concept of causation was the idea of necessary connexion (where 'that relation is of much greater importance, than any of the other two above-mention'd [—viz., *contiguity* and *succession*]' (1978, 77). In fact, Hume did not stop looking for an impression that corresponded to the idea of necessary connexion. Rather, he abandoned the route he had initially chosen, namely, the direct hunt for an impression of sensation that leads to the idea of necessary connection, and looked into what happens to the mind when it is engaged in inferences



from causes to effects in an attempt to ground the idea of necessary connexion to impressions in a roundabout way.

Brown does not for a moment deny that (B) is the way to understand Hume's project. It follows from this that Hume never took it that the idea of power (which for Hume, Brown, and all the rest is 'synonymous' to the idea of necessary connection) lacks content. For him, however, Hume's project was misguided: he looked for an impression where none can be found. He thought he had found the 'real prototype' (1822, 194) of the idea of power, but in the end he found nothing like an impression and he should have rested simply with the similarity of the of events that constitute a regularity. Nor, of course, does Brown deny that we have an idea of power. But, as we have already seen, he takes it that the idea of power arises from the belief in the future conjunction of events like the cause and events like the effect.

What then is the idea of power an idea of? On Brown's reading of Hume, Hume (a) accepts that we possess the idea of power; (b) finds its origin in an impression of reflection (the felt determination that is conditioned by the observation of constant conjunction); (c) therefore, takes it to be an idea of *something* (notably, of the transition from the idea of the cause to the idea of the effect) distinct from the sequence of events. So, Brown's Hume thinks of the transition of the mind as a third factor—something we *feel*. On this third factor, then, is the idea of power based. Indeed, Hume does *posit* a new impression—'determination'—which carries the weight of his explanation of the origin of the idea of necessary connection. Hume started with an aspect of his own positive theory, namely, that habit or custom operates on the mind to make it form a *belief* of the usual attendant of an object, and takes this *aspect* of his theory as a *datum* which will give rise to the required impression. For Hume, there is *something* that happens in the mind as a result of the observation of constant conjunction. This something is not an 'impression of sensation'. If it were, the observation of a single instance of two events following each other would have the same effect on the mind. But it does not. This something, as Stroud has nicely put it, is 'a peculiar feeling that arises from the repeated occurrence of associated perceptions'.<sup>14</sup> Hume calls it an 'internal impression, or impression of reflection' (1978, 165). In the *Enquiry*, he calls it a 'sentiment' (1974, 75). No matter what exactly it is, it must be there to ground the idea of power.

That there must be an impression corresponding to the idea of necessary connection follows from Hume's Basic Methodological Maxim. That it isn't an 'impression of sensation' follows from his analysis of what is perceived *in* the objects. That, nonetheless, something happens to the mind when a 'multiplicity of resembling instances' is observed follows from his own positive psychological theory of causal belief formation, that is, from his own account of causation as a 'natural relation'. Then, it *must* be the case that this something that happens to the mind is the sought-after impression. This something that happens to the mind is what Hume calls the feeling of 'determination'. Indeed, Hume notes: 'this determination is



the only effect of the resemblance; and therefore must be the same with power or efficacy, whose idea is derived from the resemblance' (1978, 165). Its presence in the human mind after the observation of 'resemblance in a sufficient number of instances' (1978, 165) is, as Stroud has rightly put it, 'simply a fundamental fact about human beings that Hume does not try to explain'.<sup>15</sup>

But that's precisely the problem for Brown. For him there is no third factor. Nor is there need to posit it to give content to the idea of power. There is no extra feeling of determination or whatever, but only the 'feeling of invariable antecedence' (1822, 189) that is encapsulated in a causal belief. Talk of powers, for Brown, is meant to enable us to distinguish between sequences of events that are *casual* from those that are *causal*. Powers are future-oriented. They are 'indefinitely extendable'. Belief in power is belief in future invariable sequence. No impression can afford this orientation-to-the-future corresponding to the idea of power. The impression (even the impression of past constant conjunctions) will have to be of something that has happened in the past. The idea of power cannot be 'a copy from that of which it is completely different' (1822, 189).

## V.2 The Customary Transition

Hume, as is well-known, laid emphasis on the role of habit or custom, of which he said that it is 'the great guide of human life' (1974, 44) in the formation of causal beliefs. The basic psychological inferential procedure by which the observed past co-occurrence of *Cs* and *Es* leads us to conclude (and to form the belief) that upon the fresh perception of a *C*, an *E* will (or *must*) follow is based on 'a new relation betwixt cause and effect', namely, constant conjunction. This 'new relation' is a relation among sequences of events. It says: 'like objects have always been plac'd in like relations of contiguity and succession' (1978, 88). Hume, of course, does not identify the necessary connection with the constant conjunction. The observation (or memory) of a constant conjunction generates *no* new impression in the objects perceived. The mere *multiplication* of sequences of tokens of *C* being followed by tokens of *E* adds no new impressions to those we have had from observing a single sequence. Yet, constant conjunction is the source of the inference (better: the customary transition) we make from causes to effects.

Brown is very unhappy with this way of explaining the origin of the idea of power because it gives the wrong description of the psychological mechanism by which causal beliefs arise. There is no evidence that there is a moment in time in a person's psychological life in which the observation and memory of constant conjunction generates a new belief that a sequence of events is causal. In other words, there is no moment before which some similar sequences of events are considered casual and after which are believed to be causal (cf. 1822, 157).



Nor can custom help us justify the belief that some sequences are causal while others are not simply because custom is past-oriented while causal beliefs are future-oriented. Take as evidence that *E* has succeeded *C* once and consider the belief that *E* will follow *C* forever. This belief is no less justified than if it were based on evidence that *E* has succeeded *C* one thousand times. Hume himself has shown that past repetition does not give us *reason* to expect that the future will resemble the past. Brown, of course, is not a sceptic. But that is precisely his point, namely, that scepticism cannot be avoided by an appeal to custom. Take the belief in a principle of uniformity of nature. If we could offer reasons for this belief, as we cannot, an appeal to custom would be unnecessary. If, as is the case, we cannot offer reasons, an appeal to custom is powerless, simply because what we are concerned with is the reason for the part of the belief that is future-oriented (viz., the future uniformity of nature) and custom can only tell us something about what has already happened. As we have seen, Brown's own way out is to deny that belief in uniformity (or similarity) is based on reasons, while at the same time denying that this would make it unjustified.

### V.3 A New Role for Experience

What then is the role of experience (and of the observation of repetition) in the formation of causal beliefs? To answer this, we need to bear in mind Brown's own positive view about belief in causation. Brown, you might recall, was adamant that there is no more to causation than regularity of sequence; and in particular that there is no regularity enforcer: a power, or an intermediate tie, or whatever that enforces that the causal sequence is invariable, or that it ensures its future extendability. Hence, to say that *c* causes *e* is to make a general claim about the invariable succession between events like *C* and events like *E*. Causal claims (or causal beliefs) are future-oriented. However, though for Brown causation *is* regularity, belief in causation is *not* grounded in experience of repetitions, customary transitions, and the like. Belief in causation is neither perceptual nor based on reasoning. It is what Brown called 'an intuitive belief'. As we have already seen, the key feature of an intuitive belief is that it does not rely on argument and yet it is impossible for us to disbelieve it. So, when we believe that an event caused another we believe that their relation is permanent, namely, that when similar circumstances arise, the same cause will be followed by the same effect. But this belief is intuitive. If to believe that *c* causes *e* does not require observation of repetition, what is the role of experience? Brown gave experience a prominent role, but it is more complicated than is normally assumed. To put the point briefly, experience is indispensable in finding out what regularities underlie and make true causal claims. Let me expand on it.

It is part of our constitution to have the mental tendency to attribute causes to any change in the regular pattern of things and to regard as the



cause of the change ‘the circumstances that preceded it’ (1822, 160). But being the immediate antecedent of the change is not sufficient for causation. What is necessary is this immediate antecedent being *invariable*. So it is not the case that *c* causes *e* if *c* is the immediate antecedent of *e*. Some regularity should hold between events like *c* and like *e*. Still, we believe that *c* causes *e* if we observe that *c* is the immediate antecedent of *e*—without the need to see the sequence being repeated. This is based on ‘an irresistible intuition’ (1822, 168), which carries with it the belief that *c* and *e* are invariably associated. Experience then has two roles to play. The primary role is to weaken the belief that *c* causes *e* simply because experience teaches us either that there has been a lot of irregularity in nature or that the events under consideration may be classed under very different and complex patterns. The secondary role of experience is to help us find out the actual regularities that underpin a certain sequence of events—that is, to break down complex trains of events into their constituents and find out the types of events to which an instance this causal sequence is. Here is exactly where observation of repetition and experimentation plays a crucial role. As he (1822, 160) put it, their role is

to enable us to fix with precision, where there are many antecedents and many consequents, the order in which they are to be reciprocally paired.

Here is an example Brown employs. Suppose we mix two substances for the first time and ‘a peculiar product appears’ (1822, 162). This is a causal sequence of events, the chemists believe; the mixture is the effect of mixing the two substances. This belief, according to Brown, is not conditioned on past experience. He disagrees with Hume’s line on this, which he takes it to imply that prior to repetition this sequence of events is as casual as the sequence that consists of the entering of a friend in the laboratory and creation of the new substance. What is the role of experimentation then? Not to establish that there is a causal connection but to establish what exactly it is. And it does that by carefully removing circumstances that—past experience has taught us—could causally influence the effect, for example, the presence of light or air, the impurities of the substances, and so on. In this example, experience functions in both of the ways mentioned by Brown: first, it makes us wary of the fact that there may be other causes in operation (since there have been in the past) and second it helps us establish (or acquire greater certainty) that they are not in operation.

#### V.4 Against the Ultimate Objection

What transpires is that Brown had a very interesting and innovative combination of the following two views.



(I) Causation, as it is in the world, is regularity (regular and invariable association between certain types of events).

(II) Causal beliefs do not require (in fact, they do not rely upon) observation or experience of invariable associations or repetitions.

(I) suggests that Brown defended the Regularity Theory of Causation. Hence, causation is not a singular relation between events. But (II) can be described as a view of singular causal belief. Belief in causation arises in observations of single instances of change; and yet it is belief in a regularity.

This kind of combination might sound odd. But it is not. To see why, let us examine Brown's answer to the chief objection levelled against a regularity theory of causation. This is Reid's: there can be sequences that are regular but not causal in that the regular and invariant antecedent is not the cause of the consequent. Brown made an extra effort to neutralise this objection. There is a general point to be made first. Brown (1822, 76) took as one of the sources of illusion about what causation the thought that something of distinct metaphysical sort must be added to (or must be present in) an invariable sequence of events for it to be *causal*. This, he thought, stems from the common fact that we rightly want to distinguish between something causally following something else and something *casually* following something else. We then take casual sequence to be a *mere* sequence: a relation of priority and succession between two events. We are tempted into thinking that a *causal* sequence must not be a species of sequence but something totally distinct—something that leaves no possibility open that a casual sequence might be mistaken for causation. But *cause* and *sequence* are not opposed to each other; they are similar: causation is a species of sequence, namely, invariable and uniform sequence.<sup>16</sup> Isn't then Reid's objection immediately forceful?

Brown (1822, 170–1) was not moved by Reid's objection because he thought that Reid's example of night causing day either does not describe a case of regular and invariable succession or, if it does, it can be fully captured by the regularity theory of causation (1822, 170–1). All depends on how exactly the event-types that are supposed to constitute the regular succession are identified. Given a 'vulgar' (that is, coarse-grained) description of the event-types that are supposed to be in a relation of invariable succession, there is no invariable succession and hence no causation. The night, understood as various degrees of *darkness*, is not invariably followed by day, understood as various degrees of *light*: 'they . . . rather appear to follow each other loosely and variously, like those irregular successions of events, which we denominate Accidental' (1822, 171). Given, on the other hand, a fine-grained description of the event-types, there is regularity and hence causation. Strictly speaking, night and day are not events—they are not even single phenomena, but series of phenomena grouped together by reference to some similarity and difference: degrees of darkness and degrees of light. If we



focus on ‘the successive *pairs* of that multitude of events, which we denominate night and day’ (1822, 170), and if, further, we take these events to be the positions of the earth in relation to the sun during its rotation around its axis, the motion of the earth immediately before the sunrise does cause the subsequent position of the earth in which the sunlight directly reaches the ground. In this way, the succession of night and day is explained by being reduced to a more complex regularity (picked by a more appropriate description of the causal *relata*). Brown was fully aware of the fact that an advocate of the Regularity Theory of Causation can claim that an invariable succession between *A* and *B* need not imply that *A* causes *B* or that *B* causes *A*, since *A* and *B* might be the effects of a common cause *C*.<sup>17</sup>

Brown turns on its head the problem raised by Reid. Precisely because regularity constitutes causation, where there is no causation there must be an explanation in terms of the *absence* of regularity; and where there is causation, some regularity must be present, though the grounding regularity need not be described in the vocabulary in which the causal claim is made. Hence, Brown identified the claim that the advocates of a Regularity View of Causation should make: the regularities that constitute causation need not be read off directly from the description of events that constitute the *relata* of a certain invariable sequence; but insofar as there is causation, there is a suitably described underlying regularity. As he nicely put it: ‘The generalisations of language are already made for us before we have ourselves begun to generalise.’ And this may well lead us ‘to suppose a physical relation in many cases where there is none, and to neglect it as often where it truly is’ (1822, note M). And as his biographer noted:

A cause is not always that which *appears* to be the immediate antecedent of a change, but that which is in nature the immediate antecedent.<sup>18</sup>

## VI. CONCLUSION

The points made in the last section suggest that Brown might well have been entertaining a view that has been made popular in the twentieth century by Donald Davidson, namely, that though all causation is nomological, the law under which a causal sequence is subsumed may not be stated in the vocabulary in which the singular causal statement is stated.<sup>19</sup> This, however, is an issue that cannot be pursued here.<sup>20</sup> The aim of this chapter has been to show that Brown’s philosophical theory of causation was intricate, deep, and interesting—that he did try (and at least partly succeeded) to carve a conceptual space between Reid and Hume and to combine the thought that causation as it is in the world is invariable sequence with the thought that belief in causality need not require observations of invariable successions. In light of this, Brown’s place in the history of the philosophy of causation needs to be favourably reassessed.<sup>21</sup>



## NOTES

1. For a relatively detailed discussion of Brown's philosophy of mind, see M. W. Landes, "Thomas Brown: An Associationist?", *The Philosophical Review*, 35 (1926): 447–64.
2. Abbreviations used: 1822: Thomas Brown, *Inquiry Into The Relation of Cause and Effect* (Andover: Flaggand Gould Printers, 1822). 1851: Thomas Brown, *Lectures on the Philosophy of the Human Mind* (19th edition) (Edinburgh: Adam & Charles Black, 1851). 1863: Thomas Reid, *Essays on the Active Powers of Man*, in *The Works of Thomas Reid*, preface, notes and supplementary dissertation by Sir William Hamilton, Vol. II, 6th edition (Edinburgh: Machachland & Stewart, 1863). 1974: David Hume, *An Enquiry Concerning Human Understanding*, in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, edited by L. A. Selby-Bigge & P. H. Nidditch (Oxford: Clarendon Press, 1974). 1978: David Hume, *A Treatise of Human Nature*, edited by L. A. Selby-Bigge & P. H. Nidditch, (Oxford: Clarendon Press, 1978).
3. The book had a fourth edition (identical with the third) in 1835.
4. John A. Mills, "Thomas Brown's Theory of Causation", *Journal of History of Philosophy*, 22 (1984): 207–27.
5. The New Hume was a causal realist, albeit a sceptical one. This interpretative line was inaugurated in the early 1970s, when John P. Wright published his book on Hume and has been reinvigorated in the work of Edward Craig and Galen Strawson. See John P. Wright, *The Sceptical Realism of David Hume* (Manchester: Manchester University Press, 1973); E. J. Craig, *The Mind of God and the Works of Man* (Oxford: Oxford University Press, 1987); and "Hume on Causality: Projectivist and Realist?", in *The New Hume Debate*, edited by Rupert Read & Kenneth A. Richman (London: Routledge, 2000), 113–21; and Galen Strawson, *The Secret Connexion* (Oxford: Clarendon Press, 1989). My view is that the Humean view was Hume's view too, give or take a bit; for the reasons why, see my *Causation and Explanation* (Chesham: Acumen, 2002), Ch. 1. For an excellent recent discussion of Hume's theory of causation, advancing a projectivist interpretation of his views, see Helen Beebe, *Hume on Causation* (Abingdon: Routledge, 2006).
6. This is suggested, but not explored, by N. Kemp Smith, *The Philosophy of David Hume* (London: Macmillan, 1941), 91, n2) and endorsed by Strawson, *The Secret Connexion*, 7, n.1.
7. To be more precise, Reid spoke of a being or a mind that has the power and the will to produce the effect. The quotations are from N. Wolterstorff, *Thomas Reid and the Story of Epistemology* (Cambridge: Cambridge University Press, 2001).
8. To be sure, he was a resemblance nominalist in that he took it that there are resemblances in nature and that searching for resemblances is a principle by means of which the minds works. This is how Welsh puts the matter on behalf of Brown: 'We perceive two or more objects—we are struck with their similarity in certain respects—we invent a common appellation to express the object that agree in exciting the same relative feeling'. See D. Welsh, *Account of the Life and Writings of the Thomas Brown, M.D.* (Edinburgh: W & C Tait, 1825), 287. This feeling of resemblance, Brown took it to be a simple (not further analysable) state of mind. For Brown's critique of realism and conceptualism, see Welsh, 268–88.
9. James B. Peterson, "The Empirical Theory of Causation", *The Philosophical Review*, 7 (1898): 43–61, quotation from 49.
10. *Account of the Life and Writings of the Thomas Brown, M.D.*, 114.



11. *Account of the Life and Writings of the Thomas Brown, M.D.*, 106.
12. *Account of the Life and Writings of the Thomas Brown, M.D.*, 106.
13. This is not to imply that Brown had an account of how these intuitive principles arise and what exactly they are. He did not.
14. B. Stroud, *Hume* (London: Routledge, 1977), 43.
15. *Hume*, 86. For my own reading of Hume, see my *Causation and Explanation*, Chapter 1. See also P. Millican, "Against the New Hume", in *The New Hume Debate*, edited by Rupert Read & Kenneth A. Richman, 2nd Revised Edition (London: Routledge, 2007).
16. One of Brown's critics, Lady Mary Shepherd, suggested that it is a mistake to represent causation as a species of sequence, since it is better viewed as transference. She claimed that cause and effect should be represented not as a conjunction but rather in the form  $A \times B = C$ , where the effect  $C$  is included in the 'mixture of the objects called cause'. (See *An Essay upon the Relation of Cause and Effect, Controverting the Doctrine of Mr. Hume Concerning the Nature of That Relation with Observations upon the Opinions of Dr Brown and Mr Lawrence Connected with the Same Subject* (London: Hookham, 1824), 141. Then, the effect *must* follow whenever the cause ( $A \times B$ ) is present, since it is included in it (in the same way in which the result of an arithmetic operation is included in it). Of course, this sense of inclusion is exactly what Brown had denied.
17. See also Welsh, *Account of the Life and Writings of the Thomas Brown, M.D.*, 109.
18. *Account of the Life and Writings of the Thomas Brown, M.D.*, 107.
19. D. Davidson, "Causal Relations", *Journal of Philosophy*, 64 (1967): 691–703. Reprinted in *Causation*, edited by E. Sosa & M. Tooley (Oxford: Oxford University Press 1993), 75–87.
20. For the details, see my "Regularity Theories" in *Oxford Handbook of Causation*, edited by Helen Beebe, Peter Menzies, and Chris Hitchcock (Oxford: Oxford University Press, 2009), 131–57.
21. An earlier version of this chapter was presented at the annual conference of the *British Society for the History of Philosophy* at the University of York in March 2008. Many thanks to the participants for useful comments, but especially to Martha Bolton and Eric Schliesser.



## 12 Causality and Causal Induction

### The Necessitarian Theory of Lady Mary Shepherd

*Martha Brandt Bolton*

A dispute over the nature of causality involving David Hume took place in Edinburgh in the first decade of the nineteenth century. It was sparked by a controversy involving professors of the University of Edinburgh and the city's Protestant ministers over the appointment of John Leslie to a university chair in mathematics. Although a distinguished mathematician, he was deemed unsuitable for endorsing Hume's account of cause and effect. Although largely a political controversy, the "Leslie affair" moved Thomas Brown to publish lectures on the philosophical merits of Hume's account.<sup>1</sup> Advocating a regularity theory of causality, Brown's tract was well received.<sup>2</sup> But it did not satisfy the philosophical mind of Lady Mary Shepherd, a privately educated young woman growing up outside of Edinburgh. She composed essays alleging lapses of logic on the parts of both Brown and Hume and contending for a necessitarian theory of causality and reason-based explanation of causal beliefs.<sup>3</sup> Although not made public, these writings seem to have formed the basis of work published in the 1820s. In 1818, an elaboration of Brown's theory of causality appeared in print.<sup>4</sup> Shepherd noted that contingent regularity accounts of cause and effect, despite being false and illogical in her view, remained popular. In 1824, with encouragement from her husband and probably some Cambridge faculty of the couple's acquaintance,<sup>5</sup> she published *An Essay upon the Relation of Cause and Effect, Controverting the Doctrine of Mr. Hume, etc.* Arguments in support of her positive account of causality and causal reasoning are interspersed with objections to the established alternatives. Hume is the primary target, but also Brown and a physiologist who adopted a methodology based on Brown's account. In the subsequent *Essay on the Academical or Sceptical Philosophy, etc.* (1827), the theory of causality is applied in criticising Berkeley's metaphysics, explicating belief in the existence of the external world, and treating other topical issues.<sup>6</sup>



Shepherd's main aim in these volumes is to expose harmful consequences of the causal theories currently in vogue; in particular, illogical reasoning on mundane questions about the means to secure ends, on moral choices, in scientific inquiry, in questions touching on the existence of a deity. According to Robert Blackey's *History of the Philosophy of Mind* (1848):

When [Lady Mary Shepherd] undertook a public refutation of these erroneous notions on cause and effect, it must be remembered it was at a time when they were most rampant, and widely spread over the northern parts of Britain in particular. Every young man who came from the Universities of Scotland, attempted to show off his subtlety and academic lore, by denying there was any real causation in the world; all was mere imagination, and a piece of gross vulgar credulity. Her Ladyship's efforts were, therefore, well-timed; and there is no doubt but their influence was decided in giving a considerable check to these illogical and dangerous opinions.<sup>7</sup>

Her work was known at Cambridge and read by a broader audience.<sup>8</sup> After her death, however, her writings were not reprinted. The present chapter concerns her positive theory of causality and closely connected theory of causal induction more than her pointed objections to Hume and others; but the Humean position is so familiar that it is sometimes useful to compare it with that of Shepherd.<sup>9</sup> In addition to explicating this material in its historical context, I want to suggest that from the present-day point of view, her doctrines may, in some respects, seem more right-headed than those that held sway in the period of philosophical thinking that followed her death.

## I. CAUSE AND EFFECT ARE SYNCHRONOUS

Shepherd contends that the relation between cause and effect is synchronous, not successive. For us, perhaps the best entry into her way of thinking is by understanding why it is wrong to suppose that cause and effect are temporally discrete objects or events and the sources of this error, as she sees it.

Shepherd, as well as those she opposes, purports to account for the causal beliefs human beings are naturally inclined to hold. For instance, in the course of our early experience, all cognitively competent persons acquire the belief that everything that begins to exist must have a productive cause. The aim is to explicate the intentional content and aetiology of this and other natural beliefs regarding cause and effect. Hume, and those under his influence, maintain that a cause can be defined, in part, as an event that precedes its effect in time. By contrast, Shepherd holds that this misrepresents both the idea of cause and the intentional content of our



sensory perceptions. Unless this error is corrected, there can be no genuine debate over the necessity of the causal relation.

This is because, as Shepherd argues, Hume is right about this: there is no necessary connection between two objects one of which is temporally subsequent to the other. Not persuaded by Hume's conceivability test for necessity and possibility, she offers a different argument. Assuming cause and effect are temporally discrete, the effect is 'in suspense' while the cause endures; it is evident that such a cause does not necessitate the existence of the effect. Any number of things might prevent the effect. Indeed, it is not merely possible that one should occur without the other, but factual.<sup>10</sup> The point might be elaborated this way: if there is any attribute of a cause by which it might be thought to determine an effect, it exists during precisely the same temporal interval as the cause itself; a nonexistent thing has no attributes. Defining a cause partly in terms of an event temporally antecedent to its effect precludes the possibility of a genuinely necessary connection between them. For a similar reason, it undermines the efficacy of causes. If the proximate cause produces its effect at all, it does so only during the time the cause endures; it cannot act efficaciously when it no longer exists.<sup>11</sup> Philosophers who maintain that a cause is essentially antecedent to the effect thereby stack the cards against belief in causal necessity and productivity. The upshot is that a pervasive human belief is condemned as false and its origin rendered obscure.<sup>12</sup>

There are two main sources of the mistaken opinion that causes essentially precede effects, according to Shepherd. One is lack of close attention to what happens when an effect occurs. As she explains, a cause, abstractly considered, is a union among several objects. Prior to the effect, certain objects which do not manifest this effect exist and subsequently join to form the union which produces the effect: 'Fire and wood must, indeed, be antecedent to combustion, but it is in the *union* of fire and wood, there exists *immediately* combustion as a new event in nature; . . . ' (PWMS 2. 284; also 1. 51). Often we observe the union-forming process: we throw a log on the fire. It is easy to mistake the process for the cause. Yet attention shows that the accomplished conjunction of fire and log is necessary and sufficient for ignition; combustion and its cause are strictly coextensive in time. Careless observation is one source of the presumption that causes are temporally antecedent to their effects. The fact that often we judge that a certain sort of cause, say, fire juxtaposed to wood, is present in a particular situation on the basis of observing the events that led up to its presence provides opportunity for another error of which Shepherd is keenly aware. That is conflating our evidence of the existence of something and the nature of that thing: 'The confusion of mind arising from considering *cause* as essentially an antecedency, instead of a concomitance, and of making no distinction between its *nature and operation*, and our *ability to detect its presence*, is transfused into all modern writers on Cause' (PWMS 2. 273, note).<sup>13</sup>



More important, though, the misguided notion of the causal relation receives theoretical support from an incorrect account of the intentional content of sense perception, as Shepherd sees it. That is to deny that we perceive things as having causal connections. We are said to perceive the temporal order of singular events and note the regularity with which objects of one sort are succeeded by objects of some other sort. Then because the succession of any number of perceived events does not constitute a causal connection, belief in the existence of causes must have a source extrinsic to perception. The only point in dispute is the nature of the source; for Hume, perception of a regular conjunction inculcates an associative habit of imagination; for Brown, perceiving a succession of events prompts the mind to form an intuitive belief in the corresponding universal regularity. To Shepherd's mind, these accounts fail to explain both the content and epistemic justification of causal beliefs, a lapse traceable to a deficient theory of the conceptual structure of our perceptions. By contrast, she maintains, we perceive events as having contemporaneous causes. Intentional content to this effect is built into sensory perceptions as a result of a complex perceptual process. As she sees it, sense perception involves sensations which are subsumed under ideas of certain relations by latent operations of reason.<sup>14</sup> The relations are 'true' in that, by and large, sensations actually stand in these relations.<sup>15</sup> As a result, we perceive certain sensations *as* effects of external causes:

The perception of external, continually existing, independent objects, is an affair of the understanding; it is a mental vision; the result of some notions previously in the mind, being mixed with each sensation as it arises, and thus enabling it to refer the sensations to certain reasonable causes, without resting merely in the contemplation of the sensations themselves. (PWMS 2. 168)

Although we are unaware of the mental operations from which these perceptions result, they can be expressed by a rudimentary argument. (We will consider this in the next section.) It is not just that we perceive events as having causes, but also that belief in the existence of these causes can be supported by reasoning.

Hume and Berkeley are cited for advocating an especially impoverished theory of perception, on which we perceive nothing but sensations in spatial and temporal relations. They recognise no faculty of reason capable of providing structure to the sensory given, and thus capable of representing external objects and causal connections.<sup>16</sup> As she sees it:

Idiotcy appears to be little else, than an incapacity for further perception than what resides in the immediate impressions created by the use of the five organs of sense, and the power of motion . . . and although I must allow that [children] do not, cannot argue formally on



the subject; yet, . . . their understandings take notice of, (i.e. their latent powers of observation enable them to perceive,) certain simple relations in those ideas of sensation, which are determined to their minds by the organs of sense. (PWMS 2.315)

Hume's account of the intentional content of perception, unadulterated by imagination, forces him to characterise the relation of cause and effect in terms of temporal order.<sup>17</sup> Philosophers like Reid and Brown, who have less meagre views of our cognitive faculties, still fall in line with the misguided Humean doctrines. This is the second, and more theoretically important, source of the erroneous notion of the causal relation.<sup>18</sup>

Shepherd makes a further point against this complex of errors of Hume. They force him to identify causes with sensible qualities, colours, figures, and the like. But of course we do not suppose the aroma, taste, and feel of bread cause nourishing in us, but rather think that these sensations as well as its nourishing effect are caused by the inner constitution of bread.<sup>19</sup>

## II. CAUSATION IS A COMPOSITIONAL DETERMINATION RELATION

Shepherd maintains that an external object produces effects in virtue of its constitution, a union of several constituents. Her terminology varies: causes are called 'particles', 'masses of particles', 'masses', 'natures', and in some contexts 'objects'. She defines 'object' as a kind of nature signified by a name which is partially defined by a list of its sensible qualities; for example, wood, fire, and snow are objects.<sup>20</sup> The abstract terminology used to refer to external entities more generally reflects her strict scepticism with regard to the intrinsic character of causes. They are known to us only by means of their effects. With regard to objects:

The really philosophical method of viewing the subject is this: that objects in relation to us, are nothing but masses of certain qualities, affecting certain of our senses; and which, when independent of our senses, are *unknown* powers or qualities of nature. These masses change their qualities by mixture with any other mass, and then the corresponding qualities determined to the sense must of course also change. (PWMS 1. 46)

Yet there is a structural identity between the sensory effects and the external world which enables us to know relational properties of causes: ' . . . the unknown causes of all our perceptions, are as the unknown qualities in algebra, which yet may be measured, valued, reasoned on by their signs; and the signs of these outward objects are the sensations they can create; . . . ' (PWMS 2. 48).<sup>21</sup> On this basis, we can say:



These masses change their qualities by mixture with any other mass, and then the corresponding qualities determined to the senses must of course also change. These changed qualities, are termed *effects*, or *consequents*; but are really no more than NEW QUALITIES arising from *new objects*, which have been formed by the *junctions of other objects* (previously formed). . . . (PWMS 1. 46–7)

The complex masses that enter into a mixture may, or may not, continue to exist thereafter; wood is destroyed in the union with fire.<sup>22</sup> But they have constituents which survive and subsequently enter into further mixtures.<sup>23</sup> Shepherd's abstract terminology for causes in general suggests there are basic noncomposite natures ('particles'), perhaps of several kinds, from which all composite natures ('masses of particles') are composed; but this is not explicitly stated.

The theory places effects in the category of affections or qualities. As such, they are ascribed to, or presuppose, subjects on which they depend in roughly the way figure or motion depends on the existence of a moving or figured body. The theory of causes is an adaptation of the Aristotelian ontology of substances and accidents with qualities as the only accidents. Qualities necessarily depend in an asymmetric way on the natures, or objects, to which they belong; objects do not depend in *this* way on anything else. Shepherd's innovation is in treating the notion of inherence as a biconditional necessary connection. That is, necessarily if a quality exists, it belongs to a certain object: 'Necessary connection of cause and effect is the obligation qualities have to inhere in their objects, and to exhibit their varieties according to the different human senses with which they come in contact' (PWMS 1. 63). And necessarily if an object exists, it has each of its several qualities: 'An Effect is the produced quality exhibited to the senses, as the essential property of natures so conjoined' (PWMS 1.63). Accordingly, it is impossible that there should be a merely contingent connection between a cause and its effects.<sup>24</sup>

Shepherd proposes the following definition:

A Cause . . . is such action of an object, as shall enable it, in conjunction with another, to form a new nature, capable of exhibiting qualities varying from those of either object in separation from the other. This is really to be the producer of new being. (PWMS 1. 63<sup>25</sup>)

Fire causes wax to melt just in case wax separated from fire is solid and necessarily wax in conjunction with fire is liquid.

Because effects begin and cease to exist, so do their causes: 'The beginning of every quality is perceived to be only a *change*, upon some objects already in existence; . . . ' (PWMS 1. 143, note). 'The junction of two or more . . . objects is wanted to every new creation of a new quality' (PWMS 1. 187). All natures undergo change at least with respect to their mutual



relations. There are two aspects of this that call for comment. One is the fact that natures change: they mix, separate, mutually interfere, and interact. Although this does not call for a causal explanation, on Shepherd's account, some sort of explanation would seem to be needed. Perhaps Shepherd's agnosticism about the intrinsic character of causes relieves her of the obligation to offer an account. But, I might suggest, it is open to her to say that changes in composite objects are effects of synchronous changes among their components, and changes among noncomposite natures are modelled on the following. It is allowed that a body in motion continues to move in the same direction at the same speed unless acted upon by a force, even though continued motion implies continued change in the body's relations to places and times. In a similar vein, Shepherd might say that there are no changes intrinsic to the basic enduring natures; all coexisting basic natures jointly determine (are sufficient for) changes in their mutual relations. On this account, all changes in the external world are, at bottom, changes with regard to extrinsic relations among basic 'particles'.

But even if noncomposite natures undergo no intrinsic change, their unions begin and cease to exist. As we will see, Shepherd maintains that it is necessary that a thing that begins to exist has a cause. It would seem to follow that if a compound object begins to exist, it must have a cause. Now assuming the cause is an entity individually distinct from the compound, an infinite regress of causes threatens; and if the assumption is denied, the cause of the compound is, trivially, the compound itself. Shepherd does not explicitly discuss this matter, but her theory plainly applies only to qualities, that is, entities that inhere in something which does not inhere in anything else. Accordingly, we can say this much. If the relation between a compound and its constituents is not inherence, but rather, perhaps, mereological constitution, then compounds are outside the scope of Shepherd's theory of causes. On the other hand, if a compound is correctly construed as a quality that inheres in its parts, then it has a cause, namely, its parts.

### III. TWO GENERAL NECESSARY CAUSAL PRINCIPLES

The opposition between this theory and more established accounts of causality is captured by the two causal principles for which Shepherd contends: (1) Necessarily anything that begins to exist has a cause; (2) Necessarily similar causes have similar effects in similar circumstances. It is a contradiction to deny either of these general principles.

Her case for (1) depends largely on the thesis that things that begin to exist are in the category of qualities.<sup>26</sup> As we saw, Shepherd explicates the metaphysically dependent status of qualities and the inherence relation in such a way that it is impossible that a quality exists without a temporally coextensive cause. This is the thrust of the following argument:



. . . it is a contraction to suppose things to BEGIN of themselves; for this idea is occasioned by the impression, (the observation,) that the *beginning* of every thing is but a change of that which is already in existence, and so is not the same idea, (the same quality,) as the *beginning of being*, which is independent of previous being and its changes. The two ideas are therefore *contrary* to each other . . . Changes therefore require beings already in existence, of which they are the affections or qualities; . . . The mind therefore taking notice of changes, refers them to objects of which they are the qualities. (PWMS 2. 170–1)

There is nothing special about a quality's beginning to exist as opposed to its existing at all. The contradiction is just this:<sup>27</sup> to suppose something begins to exist without a cause is to suppose an essential property exists and the object to which it is essential does not exist. This has implications regarding the formation of singular causal beliefs. Shepherd supposes we discover the causes of many particular effects by observation and a simple argument, often implicit. We observe that a particular thing begins to exist, latently reason that it must have a cause (as just explained), observe that there is only one object absent before it began to exist and present when it appeared, and conclude that this object is the cause.<sup>28</sup> The following example is hers: you are looking out the window, close your eyes, and find that the view disappears; there must be a cause of its disappearance and the fact that your eyes were open while the view persisted and closed when it desisted convinces you that closing your eyelids was the cause. Now Shepherd is well aware that you might be mistaken in believing that the position of your eyelids is the only factor absent before and present at the moment the scene was annihilated; indeed, this is also true of certain events in your visual system. But because these are relatively stable concomitants of the position of the eyelids, the conclusion is accurate enough for practical purposes. Again, you might wrongly believe your eyelids were closed at the moment the sight disappeared; but you are likely to be correct about such a thing. Because we often form particular causal beliefs in circumstances in which these two conditions are met, we quickly learn how to manage many everyday affairs, as Shepherd has it. The important point is that the form of the implicit argument is valid; all one need do is get the empirically evidenced premises right, or right enough.

This opposes Hume and Brown in the following respect: (a) reasoning, as well as observation, is needed to discover the relation of causality between particular things. It opposes Hume in a further respect: (b) in favourable circumstances, observation of a single instance, as opposed to many similar instances, suffices to inform us about a cause of a particular effect. This is not to deny that for some sorts of effects, and for purposes that demand high standards of accuracy, many trials may be needed to isolate the one or more factors the presence or absence of which varies with that of those effects; but the form of reasoning is the same, she suggests. She



cautions, however, that the cause thereby identified may only be *a* cause, as opposed to the combination of factors that constitutes the *complete* (efficient) cause.<sup>29</sup>

Principle (2) licenses generalisation from a particular case in which an A causes a B to all similar cases; that is, all cases of the same kind. Principle (2) is derived from (1).<sup>30</sup> Here again, the argument is first presented in a familiar informal manner available to all cognitively competent human beings. Suppose you discover that a particular thing that begins to exist is caused by such and such a particular object, and subsequently observe a thing individually different from, but similar to, the previously observed cause. You may conclude it will have a similar effect by reasoning: 'I observe . . . in this or any other [further] case, all the prevenient circumstances the same as before; for there is nothing to make a difference; and a difference cannot arise from itself' (PWMS 2. 45). The last clause is principle (1). The operative notion of resemblance is similarity in all causally relevant respects. The inference is valid: if A causes B in one instance, then any thing (relevantly) similar to A causes an effect (relevantly) similar to B.

Of course a person may be mistaken in judging that two or more individually different objects are the same in respects causally relevant to the effect in question. This does not affect the validity of the reasoning.<sup>31</sup> Shepherd discusses at some length the steps one can take to collect empirical evidence sufficient to make it reasonable to believe that a particular object belongs to a certain natural (causally relevant) kind, but this largely epistemic issue is not strictly relevant to our topic. To return to the derivation of (2), suppose we discover that the union of a particular fire and a particular log causes combustion, throw a different log on a different fire, and find that combustion does not occur. Principle (1) implies that some causally relevant factor which was present in the first instance is absent in the second one; otherwise combustion in the first case would have had no cause.

Other passages express the argument for (2) in a more abstract way that depends more plainly on the doctrine that causality is a compositional determination relation. The following version exploits an analogy between physical unions and arithmetical operations:

Cause and Effect, might be represented rather by  $A \times B = C$ , therefore C is INCLUDED in the MIXTURE OF THE OBJECTS called CAUSE. If C arises once from *the junction of any two bodies*; C must upon every other like conjunction, be the result; because there is *no alteration in the proportions of the quantities to make a difference*;—C is really included in the MIXTURE of A and B, although, to our senses, we are forced to *note down* (as it were) the SUM arising from their union, *after the observance of their coalescence*. (PWMS 1.141)

One might think of this as an application of the principle of sufficient reason: there is no difference between similar instances without a sufficient



reason. But although this is correct, sufficient reason applies in the domain of cause, effect, or things that begin to exist, only because causes metaphysically determine effects. The principle of sufficient reason is not a premise from which it is supposed to follow that everything that begins to exist has a cause; on the contrary, the applicability of sufficient reason follows from principle (1). To put it differently, causal generalisation is metaphysically underwritten by the necessary relation between cause and effect in general.

Shepherd maintains that general causal beliefs are formed by ‘reasoning upon an experiment’. Truth-conducive operations of reason, expressed by principle (2), applied to an empirically evidenced particular causal belief account for both the generation and justification of general causal beliefs. In her view, this compares favourably with the relatively unexplanatory theories widely admired at the time.<sup>32</sup>

It follows from (1) and (2) that it is impossible that the course of nature should change, as Shepherd does not fail to point out. She notes that Hume floats two different accounts of a change in the course of nature. According to the first, things in all other respects alike might have effects of one sort at one time and effects of a different sort at another time. This is inconsistent with (2). Shepherd does not deny that a substance in other sensible respects like snow, but salty in taste, might fall from the clouds, but she contends it would not be snow. Either it would not conform to our definition of snow (in case this is inconsistent with saline taste) or it would not have the effects that all things of the same kind as those we call ‘snow’ must have as a condition of being in that kind.<sup>33</sup> Moreover, to conceive that snow might have different effects in the future than it has had in the past, we would need to conceive the causally operative inner constitution of snow, a conceptual feat we are in no position to perform. Again, if the sun fails to rise tomorrow, which we may easily imagine, that will not show that the course of nature has changed, unless the circumstances in which the sun has risen before are exactly the same tomorrow; this is not something we can imagine in all requisite detail.<sup>34</sup>

According to Hume’s second scenario, the ‘secret nature of objects’ might change without a ‘change in the sensible qualities’ (PWMS 1. 74–75). This would conflict with (1). This principle implies that a change in the secret nature of objects must have a cause, from which it follows that the supposed change is not a change:

... if “we imagine the course of nature *may change*,” it must be under the notion of a *cause equivalent to it*: ... But for nature otherwise to change, and to vary either her “*Effects*”, or “*Secret powers*,” without varying the causes or prevenient circumstances whose junction formed the objects, whence these result;—is so obviously impossible, that we cannot even suppose the will and power of the Deity to be able to work the *contradiction*. (PWMS 1. 72<sup>35</sup>)



Miracles may be possible, Shepherd grants, but: ‘No miracle could form an *uncaused* change in nature . . . ’ (PWMS 1. 79). A miraculous interference of God might alter the course of nature, but the alteration would, thereby, have a cause. By similar reasoning, the *temporal order* of many kinds of events cannot differ in different epochs. If placing dry logs in roaring fires in the presence of oxygen, and so on, ceased to be followed by combustion at some future time, the change would require a cause; the presence of the cause would be a respect in which past and future mixtures of logs and fires were not alike. The future would, then, mark no change in the course of nature, but only a novel conjunction of causes. The sequence of events is thus bound by causal necessity even though causes are temporally coextensive with their effects. Shepherd accordingly distinguishes two sorts of necessary connections. The following makes the point specifically with regard to regularities that are observed by us:

The necessary connection therefore of cause and effect, arises from the obligation, that like qualities should arise from the junction, separation, admixture, &c. of like aggregates of external qualities. But the necessary connection of *invariable antecedency and subsequency* of successive aggregates of sensible qualities, arises from the necessity there is, that there should be invariable *sequences of effects*, when one *common cause* (or exterior object) mixes successively with different organs of sense, or various parts of the human frame, &c.

Of this obvious and important distinction, between these *two kinds of necessary connection*, the authors alluded to take no notice. (PWMS 2. 131)

Necessary regularities, by contrast with contingent ones, track the histories of objects that unite together to form other objects; their qualities must occur in the temporal order of the corresponding changes among objects. A change in the order of qualities would indicate that different sorts of objects are being tracked.<sup>36</sup> In a recent article on Hume, Simon Blackburn contrasts a world in which there are particular necessary connections between particular pairs of events to a world in which the connections between similar events are immutable. In the former, it is contingent that patterns that have occurred in the past continue to be repeated. As we will see, Shepherd makes the same point. Blackburn suggests this induces a sense of vertigo which we seek to relieve by belief in an immutable source of order: ‘a Straitjacket on the possible course of nature: something whose existence at one time guarantees constancy at any later time’.<sup>37</sup> Shepherd’s theory provides just such an immutable structure. Some vertigo may remain nonetheless, because the underlying stability is entirely consistent with radical changes ‘as far as all outward appearances and modes of detection can go’ (PWMS 2. 274).



## IV. MATHEMATICAL INDUCTION

Shepherd very naturally extends her account of causality to some parts of mathematics. As she sees it, arithmetical operations such as addition, subtraction, multiplication, division, powers, and so on, are ways of constituting distinct numerical values from certain given values. The construction of geometrical diagrams is regarded as the arrangement of lines, angles, and so on, in a region of space. Mathematical compounds are analogous to physical compound natures in that their particular constitutions determine their particular essential properties.<sup>38</sup>

As a child learning geometry, the text reports, she initially doubted that the properties present in a particular diagram would be found in another similar diagram. The reservation was resolved once she realised that ‘... the affections of *all*, were INVOLVED in *one* of *each kind*; as there was nothing to occasion a *difference* amidst their relations’ (PWMS 1. 92). That is, the fact that a particular diagram consists of such and such lines and angles in a certain relation dictates the properties the diagram exhibits: ‘[O]bjects are what their formations render them.’ Accordingly:

*... when objects are formed the same upon one occasion as another, their qualities, properties, and effects, will be similar.* It is this proposition on which mathematical demonstration, and physical induction equally, and only, rest their truth. (PWS 2. 279)

Generalisation in mathematics is, then, analogous to reasoning about nature:

... the demonstration, by means of reasoning on a diagram, is but the “*one simple and judicious experiment*,” which proves the relations of every other formed after a similar fashion. ... Could these maxims of causation be altered; could qualities begin of themselves, could (therefore) like cause produce other than like effect, all the axioms, diagrams, and demonstrations might stand as they do on the books of Euclid, without any avail as to their application to other diagrams of a similar kind and their properties: ... (PWMS 2. 285–6)

Again: ‘... one experience shows that ten taken ten times over, yields a hundred; but it is reason which proves that this result coalesces in and with its cause, and that in every step of its progress: and *that* if it once coalesces it must necessarily and invariably do so always’ (PWMS 2. 295).

Shepherd plainly regards this mode of inductive reasoning as not only perspicuous, but also a main principle of arithmetical and geometrical proof.<sup>39</sup> It is not clear to me whether she supposes that all general truths of arithmetic and geometry, let alone mathematics more generally, can be demonstrated



in this way. But she goes so far as to say: ‘the science of mathematics is truly but one branch of physics: for all conclusions its method of induction demonstrates, depend for their truth upon the implied proposition, “That like causes must have like effect,” . . . ’ (PWMS 2. 278–9).

The form of argument is familiar: necessarily a particular S is (causes) P, where ‘S’ stands for the name of a natural kind; so necessarily every S is P. The validity of the form of reasoning is grounded in Shepherd’s essentialist metaphysics. But an instantiation of the argument form replaces ‘S’ with the name of some kind, for example, ‘triangle’, ‘square of 10’. The probative force of the resulting argument depends partly on (i) whether this is in fact the name of a causally relevant kind; and (ii) if it is, whether it stands for exactly the kind of essence that determines property P. To take simple examples, (i) is false if ‘S’ is ‘figure drawn by Mary Shepherd’; (ii) is false if P is replaced by ‘has equal angles’ and S is replaced by ‘triangle’, as opposed to ‘isosceles triangle’. These considerations point to epistemic issues like those mentioned earlier in connection with physical causality. In Shepherd’s view, this sort of challenge is easier to meet in mathematics than the natural sciences:

. . . the only difficulty with respect to physical causes is the want of ability to detect their similarity with any previously given data; . . . whilst the necessity of the connexion between mathematical results and their data only arises from our being able ourselves to limit the data, which means we know that nothing can occur to make a difference in the formation of future mathematical diagrams; and which formations are the true causes for all relative proportions contained in them. (“Lady Mary Shepherd’s Metaphysics”, PWMS 1. 707<sup>40</sup>)

Mathematical theorems are more certain than causal generalisations, according to Shepherd, because we have better epistemic access to causally relevant constitutions in the former case.

In nature, we have to find sensible criteria for the presence of causally efficacious objects of the same kind. But we construct mathematical objects by known methods applied to know elements:

In the mathematics, diagrams are formed by ourselves, and we may therefore be always *sure* of our future and universal conclusions; because we frame an *hypothesis*, and examine by one experiment, (i.e. one experience,) the relations which arise; and the same data being given to all future ages, there is nothing supposed which can make any difference amidst the relations; . . . (PWMS 2. 288)

Now this may suffice to insure the certainty of inductions on the entirely determinate sorts of objects formed by particular mathematical operations, for example, inductions on the kind, isosceles triangle with altitude 3 cm.



But most theorems concern less determinate kinds. A particular isosceles triangle has an area equal to half the product of its altitude and base in virtue of only a certain aspect of its constitution, which may be more difficult to identify than Shepherd seems to realise. Goldbach's Conjecture may illustrate the point.<sup>41</sup> Even though we can specify a kind of constitution common to all even numbers, being the product of 2 and another whole number, this does not suffice to prove that all such numbers are equal to the sum of two prime numbers. This is not a completely devastating objection to Shepherd's theory, but it points to the fact that she offers no principled way to identify which of the several kinds instantiated by the constitution of a particular figure is causally responsible for its various properties. Perhaps she does not fully grasp the challenge, but much the same is true of other philosophers of her time. Several of them advocate mathematical proof by generalising from particular, or singular, instances; consider the otherwise very dissimilar accounts of Hume and Kant. It is doubtful that either of them resolves the problem of identifying which of the more and less determinate kinds to which the base instance belongs supports a specific generalisation.<sup>42</sup>

Viewed in historical context, Shepherd's account of the metaphysical and epistemic ground of mathematics has a unique place. Most philosophers of the period agree that mathematical truths are universal and necessary and for that reason *a priori*. But if mathematical truths are proved *a priori* from self-evident axioms or definitions, it is difficult to make out that the fit between mathematical truths and relations among measurable physical magnitudes is more than a contingent regularity.<sup>43</sup> We might put it this way: a general necessary truth imposes necessity on an actual particular thing only if this thing is an instance of the general truth; but how are we to know that it is, if we have no empirical knowledge of necessary truths about actual particular things? It is well known that Kant intends transcendental idealism to bridge the gap between necessary *a priori* propositions, including mathematical theorems, and the realm of experience. While Shepherd expresses little interest in this problem, her theory of mathematical induction offers a solution.

Dugald Stewart is Shepherd's foil in an essay included in her second book.<sup>44</sup> Stewart promotes a familiar reason for separating the method and results of natural science from those of mathematics. We cannot come to know laws of physics without experience; but observation of particular bodies is of limited justificatory force. According to Stewart, enumerative generalisation is licensed by a 'law of thought' to the effect that the results of a particular experiment will hold in the future. The law conjoined with empirical reports of particular instances justifies belief of the physical generalisation, but not its necessity. By contrast, Stewart maintains, the universal truths of mathematics are necessary; the evidence that supports mathematical truths must be of a radically different sort than empirical reports and the 'law of thought' by which causal laws are established. He



concludes that mathematical and physical induction have different logical structures, the former demonstrative, the latter not.<sup>45</sup> As we might expect, Shepherd traces both lines of thought to a mistaken account of causality. Failing to understand that physical causal generalisations have metaphysical ground in the necessary connection between the constitutions of particular objects and their sensible qualities, Stewart fails to see that general mathematical truths are proved on the basis of observed properties of particular mathematical constructions and principle (2).

A view on which we have *a posteriori* knowledge of necessary truths is of some interest nowadays. It is worth taking a closer look at why Shepherd subscribes to it. It is based largely her theory that certain necessary truths are built into the intentional content of sensory perceptions of particular things. That is, we perceive qualities as effects which necessarily inhere in and are causally determined by external objects. We naturally credit the intentional content of observations, provided they are mutually consistent. Suppose the world is, by and large, as we take it to be in this regard; epistemology is not our topic. Shepherd is then in position to say that if a person observes a particular thing, she correctly takes it that this thing causally, necessarily determines her sensations, so she has empirical cognition of a singular necessary truth.

To her mind, the crux of the issue regarding the nature of causality is the human ability to form abstract ideas. In general: 'The *faculty of abstraction* is truly the origin of science. By abstraction, is meant the consideration of any quality apart from others with which it may usually be united, in order to notice what inferences may be drawn from its nature' (PWMS 2. 291). It is well known that Hume denies that we can form ideas of things we perceive in abstraction from all sensory aspects of those things. But Shepherd contends we need to frame just such an abstract conception to reason correctly about the relation of cause and effect.<sup>46</sup> Hume's difficulty finding the source of the idea of necessary connection is traced to his repudiation of abstract ideas of the requisite sort. It reduces him to seeking necessity in the sensory aspects of particular experiences, where of course it cannot be found.<sup>47</sup> This, in turn, leads to the error of logic which was mentioned earlier:

Indeed, Mr. Hume makes a great mistake in supposing it is necessary to demonstrate, in every particular instance, what *particular* Effect must necessarily flow from its object, in order to gain the idea of *necessary Connexion*. The *how* and the *why* have nothing to do with the general reasoning affecting the general proposition; for 'whether like Causes shall produce like Effects' is *not* a question exactly the same as whether 'such particular causes shall have such *particular* effects'? Which Mr. Hume seems to consider as precisely of the same import. (PWMS 1. 59)

Even if we had an impression of necessary connection between two particular objects, we would not thereby discover a necessary connection between



all otherwise similar pairs of objects, let alone cause and effect in general. Notice that the same logical point pertains to knowledge of general necessary truths about numbers proportions and the like.<sup>48</sup> The necessity must be perceived anew in each particular instance.

Still, if we had sense impressions of causal necessity of the sort Hume apparently seeks, then we would be presented with a necessary bond between two particular objects; we would understand why and how it is forged. This is more than Shepherd can claim. For her, we establish necessity by considering the *abstract* cause-effect relation, and *thereby* prove it in general. This leaves us without insight into the nature or workings of this or that specific necessity. To be sure, causal necessity is founded on the essences of things, but we neither know what these essences are nor how they are connected with essential properties. Notice that this is no less true of mathematical necessity than physical causality, on her account. Shepherd's opponents seem to think that *a priori* necessity is transparent, but more recent philosophers find the nature of mathematical necessity elusive. Finally, although Reid, Stewart, and Brown do not share Hume's qualms about abstract ideas, neither do they agree with Shepherd that the metaphysical profile of a cause is included in the conceptual structure of sense perception. Shepherd has a firm theoretical grip on the wedge this drives between the metaphysical notion of necessity and the notion of *a priori* necessities which are supposedly revealed as necessary in the world. Hume's doctrine that necessary truths are 'relations of ideas' conflates the evidence of necessity with necessity and, in addition, renders necessary truths entirely irrelevant to anything other than ideas. If Stewart and the others do not follow Hume on the former point, they do on the latter.<sup>49</sup>

## NOTES

1. Thomas Brown, *Observations on the Nature and Tendency of the Doctrine of Mr. Hume Concerning the Relation of Cause and Effect* (Edinburgh: Mundell and Son, 1805). See Stathis Psillos, "Regularities All the Way Down: Thomas Brown's Philosophy of Causation", included in this volume; also John A. Mills, "Thomas Brown's Theory of Causation", *Journal of the History of Philosophy*, 22 (1984): 207–27.
2. Leslie was appointed to the chair and Hume's theory of causality became increasingly influential; see Mills, "Thomas Brown's Theory of Causation", 209, the David Walsh's "Memoir of Dr. Brown" in the introduction to Thomas Brown, *Inquiry into the Relation of Cause and Effect* (London: H. G. Bohn, 1835), facs. edn. with introduction by Bernard E. Rollin (Delmar, NY: Scholar's Facsimiles & Reprints, 1977); also see Introduction to *Philosophical Works of Lady Mary Shepherd*, edited by Jennifer McRobert, 2 vols. (Bristol: Thoemmes Press, 2000). Hereafter abbreviated PWMS.
3. On Shepherd's life, see DNB, vol. 50, and the Introduction, PWMS.
4. There is evidence that at roughly this time, Shepherd wrote a treatise on causality which cannot be positively identified now. McRobert tentatively identifies it with a book published anonymously in 1819; a large part of it is



- an extended argument against Brown's account of the relation between cause and effect. A facsimile edition is included in PWMS, 1.
5. The Cambridge philosopher Robert Whewell, who later produced an important theory of causal induction, is reported to have used one of Shepherd's books as a text. The couple were lifelong correspondents of another member of the Cambridge faculty, Charles Babbage; see DNB.
  6. PWMS contains facsimile reproductions of all of Shepherd's publications. On the critique of Berkeley, see Margaret Atherton, "Lady Mary Shepherd's Case against George Berkeley", *British Journal for the History of Philosophy*, 4 (1996), 347–66.
  7. Robert Blackey, *History of the Philosophy of Mind*, 4 vols. (London: Trelawney Wm. Saunders, 1848), v. 4, 43.
  8. There was a published exchange between Shepherd and John Fearn, a retired officer in the navy; see PWMS, 1.
  9. On Shepherd's critique of Hume, see Martha Brandt Bolton, "Lady Mary Shepherd and David Hume on Cause and Effect", in *Feminist History of Philosophy*, edited by Eileen O'Neill and Marcy Lascano (Dordrecht, Netherlands: Springer, expected 2010).
  10. Hume's whole argument consists in the possibility of imagining an effect 'non-existent this minute,' and 'existing the next'; and . . . has no other way of supporting his . . . notion of the beginning of existence by itself, except under the idea of an effect in suspense; which is still a relative term, and begs the question for the necessity of its correlative, i.e. its cause . . . ' (PWMS 1. 38). Shepherd traces Brown's view that there is a logical gap between what has been observed so far and what will happen in future to the same incorrect assumption about the temporal priority of cause to effect; see PWMS 1. 141.
  11. Hume says that causes are 'only the forerunners' of effects 'and hence he infers, that there is no PRODUCTIVE PRINCIPLE, that there is only antecedency and subsequence of events of an arbitrary kind; . . . ' (PWMS 2. 297).
  12. PWMS 1. 129
  13. Shepherd spots the same confusion in passages where Hume takes the fact that we do not know that the same *sensible qualities* always correspond to the same secret powers to show that we have no reason to believe that similar *causes* will have similar effects; see *Treatise*, 63–64; PWMS 1. 60.
  14. 'Thus simple sensation has many varieties or kinds. When it refers to no other existence than itself, it should be considered as *sensation* properly and immediately. In this sense we have the *sensation of an idea*; but then *idea* refers to an existence always considered independent of sensation; which idea is only its *sign, representative*, IMAGE, or whatever name it may please philosophy to term it' (PWMS 2. 136).
  15. PWMS 2. 267. Shepherd's account of our justification for believing that sensations typically are as we perceive them to be in this respect, i.e., they actually are related to external causes, etc., is not the topic of this chapter. Roughly speaking, she holds what we might call an evidence-based fallibilist theory of justification: 'And as the knowledge of external nature is but an inference from reason, either from the relation of cause and effect, analogies, probabilities, &c. so its absolute independency of each mind, can have no further certainty than such inference, however strong it may be, can afford' (PWMS 2. 172–3).
  16. 'I conceive it impossible to have a complete conviction that every Effect is inherent, or contained in its Cause, until the mind be imbued with the knowledge, that objects are but unknown circumstances in Nature, when



- unperceived by the senses; which when perceived, exhibit their appropriate qualities accordingly; . . . ’ (PWMS 1. 42, note).
17. E.g., PWMS 1.129–30.
  18. ‘Thus I consider it to be the want of separating our *perceptions* from their causes, which has given occasion to the false notion, viz. that of the *successive effects perceived*, the *antecedent* are causes and the *subsequent* effects’ (PWMS 2. 306).
  19. ‘No child or ignorant person supposes that it is the motion, figure, brilliancy, or colour of fire (when separated from the outward permanent causes of those qualities,) which efficiently governs the burning of flesh; for that these antecedent qualities *after* being determined upon the mind, *are* the only *causes of any subsequent burning*, is a *discovery* which they leave to philosophers to make; . . . ’ (PWMS 22. 317–8; also 125–30).
  20. The definition is established by convention, but the objects that satisfy the definition typically have the same causally relevant nature; see PWMS 1. 157–8.
  21. Bertrand Russell, *Human Knowledge: Its Scope and Limits* (London: George Allen and Unwin, 1948), 256, propounds a similar view of the extent of our knowledge of the external world .
  22. ‘ . . . in general, an entire mixture, junction, and concussion of qualities, involves the whole original objects in ruins, whilst it strikes out a vast many new and altered ones, creating other masses, other complex objects, totally unlike those whose union was their Cause. On the other hand, it sometimes appears that nature intends to render one individual essence, the prime object intended to be preserved, . . . as in the growth of plants and animals; or vigour, improvement, character, individuality, &c. of the sentient principle [sc. mind].’ (PWMS 1.189; also see 2. 172–3)
  23. This might be suggested by passages such as this: ‘Now it is the *formation* of the particles, (whatever particles may be,) which renders exterior objects what they are, and of any certain definite constitution; . . . ’ (PWMS 2. 304).
  24. With reference to a popular view that qualities inhere in objects in virtue of contingent laws of nature: ‘No arbitrary law can create a mutual interference of qualities. Indeed, I have in vain endeavoured to find what philosophers exactly mean by the word law; the only rational signification is that mode of being, or action, or relation of qualities, which as Mr. Locke says, “renders an essence that which is it and not another.” But it appears to me, as though they mean it to signify an arbitrary rule which matter would observe without there being a necessity for it in any physical cause. This is impossible’ (PWMS 2. 313, note).
  25. Similarly, ‘Power is but another word for efficient cause, or “productive principle;” and signifies the *property* which lies in the *secret nature of the objects*, when unobserved by the senses, and which determines the qualities that can be exhibited to them upon every new conjunction’ (PWMS 1. 63–4).
  26. ‘ . . . the mind of man was forced to look upon all things which begin to exist as *dependent* QUALITIES; and thus, that an object could neither depend upon *itself for existence*, nor yet depend upon *nothing*’ (PWMS 1. 37; also 86).
  27. One passage puts it this way: ‘ . . . if any particular quality were supposed to *begin of itself*, the following contradiction would arise, viz. That the beginning of existence, which is a quality of being, could belong to a being not yet in existence; . . . ’ (PWMS 2. 290).
  28. To make the argument fully explicit: (i) Necessarily everything that begins to exist has a cause; (ii) X begins to exist; (iii) the cause of an effect is a factor present if and only if the effect exists and absent when the effect does not



exist; (iv) in this situation, Y is the only factor present when the effect exists and absent when the effect does not exist; (v) so Y is the cause of X. As J. S. Mill, *A System of Logic*, edited by J. M. Robson, in *Collected Works of John Stuart Mill*, v. 7 (Indianapolis: Liberty Fund, 1973; reprinted 2006), 391, was shortly to point out, (iii) is not the only criterion useful for the discovery of causes; moreover, a different method is needed to investigate the effects of a given object; also see J. L. Mackie, *Cement of the Universe* (Oxford: Clarendon Press, 1980; reprinted 1999), 59–87. Shepherd refines (iii) in view of the distinction between a part of the cause and the whole cause of an effect; see note immediately following.

29. ‘ . . . [A]ny *one* of the qualities or objects needful in order to the formation of another, may be termed a *Cause*, because absolutely necessary, . . . The *whole* number of objects existing, which are necessary to it [sc. production of an effect], may also, under one complex idea, be deemed *the one whole cause* necessary. . . . The *union of these*, is the proximate Cause of, and is *one* with the effect . . . This ambiguity, arising from the necessity of naming each object, wanted to an end, and al that are wanted to it, and the junction necessary to it, the Cause of it, is a fruitful source of error in every branch of analytic philosophy’ (PWMS 1. 187–8).
30. “*Like cause must exhibit like effect;*” and this axiom depends on the principle, that “No quality *can begin its own existence*” (PWMS 2. 290; the internal quotes mark material from the earlier book).
31. ‘When more trials are needed than one, it is in order to detect the circumstances, not to lay a *foundation* for the *general principle*, that LIKE Cause repeated, a LIKE Effect will take place’ (PWMS 1. 44, note).
32. ‘It is not therefore necessary to have recourse to any instinct or principle of nature, which we know nothing of, in order to explain the source of those ideas which govern their [sc. children’s] expectations’ (PWMS 2. 316). Shortly after: ‘That class of ideas which Dr. Reid terms instinctive, and Mr. D. Stewart considers as composed of simple ideas not formed by the senses, but generated upon certain fit occasions for their production, I consider to be the conclusions of a latent reasoning’ (PWMS 2. 170).
33. Shepherd’s account of the signification of names of natural kinds is similar to Locke’s account, which has both a descriptive and referential aspect, as we might put it; see Bolton, “Lady Mary Shepherd and David Hume”, 30–32.
34. ‘Mr. Hume makes also a great mistake in supposing because we can conceive in the existence of objects contrary to our experience, that therefore they may really exist in nature; for it by no means follows that things which are incongruous in nature, may not be contemplated by the imagination, and received as possible until reason shows the contrary’ (PWMS 1. 83; also 2. 322).
35. The internal quotation marks indicate views held by Hume. Also see PWMS 1. 69–70.
36. ‘Mr. Hume is so far from being correct in supposing that *regular conjunction* GENERATES *the idea of causation*, that on the contrary, *it is only itself looked upon as AN EFFECT of its own regular cause*’ (PWMS 2.237, note).
37. Simon Blackburn, “Hume and Thick Connections”, *Philosophy and Phenomenological Research*, 50 supp. (1990), 241.
38. I suggested earlier that Shepherd may suppose that the relation between a physical compound and its constitutions is that of whole to parts; there is no reason to think she inclines to that view of mathematical compounds, as far as I can see.
39. ‘All mathematical demonstration is built upon the notion; that where quantities, or diagrams, resemble each other, the relations which are true, with



respect to ONE of each kind will be true with respect to *all* others of a *like* kind; ONLY *because there is nothing else to make a difference amongst them*' (PWMS 1. 77).

40. '... the reason why some other branches of science are less secure [sc. than mathematics] in their conclusions, is merely because of the difficulty there in *tracing the original FORMATIONS of the objects*, without inpugning [*sic*] in the smallest degree, the universality and necessity ...' (PWMS 2. 280).
41. I am grateful to Jeffrey Russell for this example. According to the hypothesis, every even number greater than or equal to 4 is the sum of two prime numbers. Although no counterexample has yet been found, no one has found a proof so far.
42. This is plain with regard to Hume (following Berkeley); I have been unable to find that Kant has a solution.
43. As Lisa Shabel puts it: 'The state of modern mathematical practice called for a modern *philosopher* of mathematics to answer two interrelated questions. Given that mathematical ontology includes quantifiable *empirical* objects, how to explain the paradigmatic feature of pure mathematical reasoning: universality, certainty, necessity.' Lisa Shabel, "A Priority and Application: Philosophy of Mathematics in the Modern Period", in *Oxford Handbook of Philosophy of Mathematics and Logic*, ed. Stewart Shapiro (Oxford: Oxford University Press, 2007), 30.
44. PWMS 2. 271–95.
45. Stewart aims this point at Laplace, with whom Shepherd finds herself in agreement; see PWMS 2. 277–8.
46. PWMS 1. 34; 2. 291.
47. 'If we possess ... any idea of power in general, we must be able to perceive some particular species of it; ... we must be able to place this power in a particular being, and conceive that being as endow'd with a real force and energy, by which such a particular effect necessarily results from its operation. We must distinctly and particularly conceive the connexion betwixt the cause and effect, and be able to pronounce, from a simple view of the one, that it must be follow'd or preceded by the other. This is the true manner of conceiving a particular power in a particular body: and a general idea being impossible without an individual; whether the latter is impossible, 'tis certain the former can never exist' (*Treatise*, 109).
48. Necessary truths are classified as 'relations of ideas', said to be known on the basis of ideas alone, and unaffected by anything extrinsic to those ideas; see *Treatise*, 51, 66.
49. Versions of this material were presented to the Colby College Colloquium in 2006, the Margaret Wilson Conference, at UC San Diego in 2006, and the BSHP Conference "Causality: 1500–2000", at York University in 2008. The members of these audiences whose questions increased my understanding of Shepherd's thought are too numerous to mention individually, but I am indebted to them all.



# Contributors

**Keith Allen** is Lecturer in Philosophy at the University of York. He is the author of a number of articles on issues relating to colour, perception, and early modern philosophy, particularly Locke.

**Martha Brandt Bolton** is Professor of Philosophy at Rutgers University. She is author of articles on many figures and topics in the history of early modern philosophy. Recent articles include “The Taxonomy of Ideas in Locke’s *Essay*”, in *The Cambridge Companion to Locke’s Essay*, edited by Lex Newman (Cambridge, 2007), “Locke and Leibniz on the Structure of Substance and Powers: the Metaphysics of Moral Subjects”, *Studies on Locke*, edited by Sarah Hutton and Paul Schuurman (Springer, 2008), and “Berkeley and Mental Representation” in *New Interpretations of Berkeley’s Thought*, edited by Stephen H. Daniel (Humanity Books, 2008).

**Angela Breitenbach** has been Lecturer in Philosophy at the University of East Anglia since 2009. Before coming to UEA, she was Junior Research Fellow in Philosophy at Sidney Sussex College, Cambridge. Her research is centred around the philosophy of Kant, the philosophy of nature and the environment, and the philosophy of biology. Her recent publications include *Die Analogie von Vernunft und Natur: Eine Umweltphilosophie nach Kant* (Walter de Gruyter, 2009) [*The Analogy of Reason and Nature: Kantian Environmental Philosophy*]; ‘Teleology in Biology: A Kantian Approach’, *Kant Yearbook*, 1 (2009), 31–56; and ‘Two Views on Nature: A Solution to Kant’s Antinomy of Mechanism and Teleology’, *British Journal for the History of Philosophy*, 16 (2008), 351–369.

**William Eaton** is Assistant Professor of Philosophy at Georgia Southern University. He is the author of *Boyle on Fire: The Mechanical Revolution in Scientific Explanation* (Continuum, 2005), and co-author of “How to Malebranche an al Ghazali: Causation, Miracles and Natural Events”, *Revue Roumaine de Philosophie*, 54 (2), 2010, with Robert Higginson. His research interests include early modern philosophy, contemporary analytic metaphysics, and the history and philosophy of science.



**Boris Hennig** is currently teaching and researching at Humboldt University Berlin. He has published on Descartes' concept of consciousness, Aristotle's four causes and other topics pertaining to metaphysics, metaethics, and their history.

**Robert Higginson** is Visiting Professor of Philosophy at Georgia Southern University. He is the co-author of "How to Malebranche an al Ghazali: Causation, Miracles and Natural Events", *Revue Roumaine de Philosophie*, 54 (2), 2010, with William Eaton. His research interests include early modern philosophy, Leibniz, and contemporary analytic metaphysics.

**Peter Millican** is Gilbert Ryle Fellow in Philosophy and Reader in Early Modern Philosophy at Hertford College, University of Oxford. His philosophical interests and publications cover a wide range, from epistemology, philosophy of language and computing, to ethics and philosophy of religion, but with a particular focus on David Hume and related topics. From 2005 until 2010 he was Co-Editor of the journal *Hume Studies*, and he is currently Illumni David Hume Fellow at the Institute for Advanced Studies in the Humanities at the University of Edinburgh.

**Pauline Phemister** is Reader in Philosophy at the University of Edinburgh and Deputy Director of the Institute for Advanced Studies in the Humanities. She is author of *Leibniz and the Natural World: Activity, Passivity and Corporeal Substances in the Philosophy of Leibniz* (Springer, 2005) and *The Rationalists: Descartes, Spinoza and Leibniz* (Polity, 2006).

**Stathis Psillos** is Professor of Philosophy of Science and Metaphysics in the University of Athens, Greece. He is the author of *Knowing the Structure of Nature* (Palgrave MacMillan, 2009), *Philosophy of Science A–Z* (Edinburgh, 2007), *Causation and Explanation* (Acumen, 2002)—which was awarded the British Society for the Philosophy of Science Past President's prize—and *Scientific Realism: How Science Tracks Truth* (Routledge, 1999). He has edited (with Martin Curd) the *Routledge Companion to the Philosophy of Science* (Routledge, 2008)—which was a Choice outstanding academic title for 2008. He has published over 70 articles in learned journals and edited collections. He has served as the President of the European Philosophy of Science Association (2007–2009), and he's currently Editor of *Metascience*.

**Constantine Sandis** is Reader in Philosophy at Oxford Brookes University and NYU in London. He recently completed a monograph on *The Things We Do and Why We Do Them* (Palgrave Macmillan, 2010) and has also published on several aspects of Hume's philosophy. He is the editor of



*New Essays on the Explanation of Action* (Palgrave Macmillan, 2009), *A Companion to Philosophy of Action* (with Timothy O'Connor, Wiley-Blackwell, 2010), *Hegel on Action* (with Arto Laitinen, Palgrave Macmillan, 2010), and *Human Nature* (with Mark Cain, Cambridge, 2011).

**Eric Schliesser** is a BOF Research Professor, Philosophy and Moral Sciences, at Ghent University. He is the co-editor of *New Voices on Adam Smith* (Routledge, 2006); the co-editor of *Interpreting Newton: Critical Essays* (Cambridge, in press); the co-editor of the *Oxford Handbook on Isaac Newton* (Oxford, forthcoming). He is currently working on two monographs one, *Adam Smith*, which will appear in the prestigious Routledge Philosophers Series, and another on *The Decline and Fall of Chicago Economics*. He has published widely on early modern figures such as Hume, Newton, Berkeley, Adam Smith, Huygens, Spinoza, Sophie de Grouchy as well as on 20th century philosophy of economics.

**Tad M. Schmaltz** is Professor of Philosophy at the University of Michigan, Ann Arbor. He has published articles and book chapters on various topics in early modern philosophy, and is the author of *Malebranche's Theory of the Soul* (Oxford, 1996), *Radical Cartesianism* (Cambridge, 2002), and *Descartes on Causation* (Oxford, 2008). He is a co-editor of the *Historical Dictionary of Descartes and Cartesian Philosophy* (Scarecrow, 2003), and is the editor of *Receptions of Descartes* (Routledge, 2005). Currently he is working on two edited volumes, the first a collection of essays on the history of the concept of efficient causation, for the Oxford Philosophical Concepts series, and the second a collection of essays on problems and prospects for integrating history and philosophy of science, which he is editing with the historian of science Seymour Mauskopf for the Boston Studies in the Philosophy of Science series. He also is continuing to work on receptions of Cartesianism, the influence of late scholasticism on early modern thought, the nature of the "Scientific Revolution," and early modern accounts of substance, causation and freedom.

**Timothy Stanton** is a Lecturer in Politics at the University of York. His research interests centre on the philosophical, political, and religious thought of John Locke and its relations to (i) liberalism, (ii) the thought of Hobbes and of (iii) Locke's intellectual enemies. He has published several essays on Locke and Hobbes and their reception and interpretation. He is currently working on a critical edition of Locke's unpublished 'Defence of nonconformity' against the strictures of Edward Stillingfleet, and, with Dr Jon Parkin, has edited a collection of essays, *Natural Law and Toleration in the Early Enlightenment*, a volume of Proceedings of the British Academy (Oxford, forthcoming).



**Tom Stoneham** is Professor of Philosophy at the University of York. He is the author of *Berkeley's World* (Oxford, 2002) and several articles on philosophy of mind, metaphysics and early modern idealism.

**David Wootton** is Anniversary Professor of History at the University of York. His books include *Galileo: Watcher of the Skies* (Yale, 2010) and *Bad Medicine: Doctors Doing Harm Since Hippocrates* (Oxford, 2006). He is writing a history of the Scientific Revolution.



# Index

## A

accidents, 39–40, 46n24, 48, 88, 247  
action: bodily (*see* body-body causation; mind-body causation);  
Divine, 31–37, 89, 90; human, 80, 106–107, 117, 128–131, 134, 136–138, 152, 153–153, 168, 170, 225  
affections, 51, 74, 77–78, 247, 253  
Alhacen, 25–26  
Allison, Henry, 168–169, 218n40, 218n41, 219n43  
Anscombe, G.E.M., 193  
Aquinas, Thomas, 33, 36  
Aristotelian: explanation of heaviness, 15–16, 21; first philosophy, 104–105; forms, 6, 79–81, 109, 113; real (or occult) qualities, 3, 39–40, 145; theories of causation, 11, 18–20, 25–26, 60, 80, 108–109, 112, 113; theories of primary and secondary causes, 31–35; theory of generation, 108–110; theory of motion, 4, 21, 52; theory of predicables, 39–40  
Archimedean science, 15–17, 20, 22, 26–27  
Aristotle, 6, 19, 20–21, 25–26, 39, 104–105, 108–109, 112, 113.  
*See also* Aristotelian  
Arnauld, Antoine, 57, 61  
Arthur, Richard, 46n15  
astrology, 22, 26  
Atherton, Margaret, 258n6  
attributes, 41, 68, 74–78, 80–1, 88, 93, 110  
Ayer, A.J., 172, 186n42

## B

Babbage, Charles, 258n5  
Bacon, Francis, 27, 66, 79–80  
Baier, Annette, 187n53  
Bailey's *Universal Etymological Dictionary*, 176  
Baliani, Giovanni Battista, 14, 17, 18, 23, 29n20  
Beck, Lewis White, 218n41  
Beebee, Helen, 7, 133–142, 144, 153, 157, 158, 160n20, 200n17, 240n5  
belief: Hume's theory of, 146, 150, 167, 179; in the existence of the external world, 137–140, 169–170, 242, 259n21; intuitive or basic, 10, 220, 231–232, 236–237, 243, 245; natural, 133, 142, 170. *See also* causal belief  
Bennett, Jonathan, 81n7, 164n61  
Bergson, Henri, 45n14  
Berkeley, George, 242, 245, 261n42  
Biagioli, Mario, 13  
Biel, Gabriel, 32  
Blackburn, Simon, 158n2, 172, 252  
Blackey, Robert, 243  
Bodin, Jean, 108  
body-body causation, 3–5, 31–2, 39, 42, 48, 53–61. *See also* mind-body causation; substance, material  
body. *See* substance, material  
Bolton, Martha Brandt, 9, 10–11, 242–261  
Boulter, S., 185n20  
Boyle, Robert, 17, 18, 24, 51  
Brahe, Tycho, 14, 15



Breitenbach, Angela, 9, 201–219

Broughton, Janet, 124, 163n49,  
185n22

Brown, Stuart, 102n4

Brown, Thomas, 1, 9–10, 220–222,  
224–241, 242, 245, 246, 249,  
257

Buchdahl, Gerd, 218n41

Buckle, Stephen, 124, 160n20, 160n26,  
161n30, 185n18

## C

Carriero, John, 37, 70, 73, 79, 84nn43–  
44

causal belief: 192; Brown's theory of,  
10, 221, 224, 231–233, 235–  
238, 245; Hume's theory of, 10,  
133, 135, 137, 142, 143–144,  
150, 152, 170, 234, 234–235;  
Shepherd's theory of, 10, 242,  
243–244, 245, 249, 251; singu-  
lar, 10, 238, 239, 249, 255, 256

causal chain, 224; deviant, 194

causal nexus, 172, 182, 207, 221

causal reasoning, 11, 145, 168, 232,  
242, 249–251, 253

cause: efficient, 1, 6, 17, 20, 21, 26, 33,  
55–57, 60, 79–80, 91–93, 109,  
111, 112–113, 114, 116–117,  
207, 223, 226, 230–231, 250,  
259n25; final, 1, 6, 11, 17, 20,  
66, 80, 91–92, 93, 109, 111,  
112–113, 117, 205, 214; formal,  
6, 20, 21, 26, 79–81, 109–110,  
111, 113, 114, 116–117; imma-  
nent, 5, 78–80, 81n5; material,  
6, 20, 21, 80, 109, 11, 112,  
116–117; primary, 3, 31–38, 43,  
90, 222; secondary, 3, 6, 31–35,  
36, 38–44, 60–61, 68–70,  
89, 222; secret or hidden, 24,  
125–126, 135, 146, 148, 174,  
176, 181–185, 251; *secundum*  
*esse* and *secundum fieri*, 41–42

Chakravartty, Anjan, 189–190

cinematic view of Divine conservation,  
35, 49–51

Clarke, Samuel, 137, 139, 158

Clatterbaugh, Kenneth, 121n39

Cohen, I. Bernard, 82n16, 83n26

Collingwood, R.G. 195

collision, 32, 42, 43–44, 47n34, 57–59

Colombe, Lodovico delle, 25

commonweal, 108–11

commonwealth, the, 6, 7, 104, 105,  
106–108, 111, 114–115,  
116–119

compositional determination, 246–248,  
250, 253

conatus, 5, 65–67, 69–70, 73–74, 76,  
80–81, 93

concurrentism, 2–3, 31–35, 43

*concurus* (concurrence), 31–38, 44, 65,  
68–69, 71–73

conservation, Divine, 2–3, 4, 31–35,  
36, 37–38, 39, 43–44, 49–51,  
53, 65, 69, 72

conservationism, mere, 2–3, 31–5

constant conjunctions, 8, 10, 128–130,  
134–136, 138, 141, 145, 148,  
153–154, 157, 168, 170, 178,  
179, 180, 182, 232, 233–235,  
245. *See also* necessary connex-  
ions

Cordemoy, Géraud de, 49

Costa, Michael, 124

counterfactual definition of cause, 231

Craig, E. J., 124, 142, 240n5

creation: Divine, 2–4, 31–35, 37–38,  
49–52, 59–60, 70–72, 110, 204;  
pro-, 28, 108–110; pro-, Divine,  
70–72

Cudworth, Ralph, 175

Curley, Edwin, 82n8

## D

Dardis, Anthony, 196–197

Davidson, Donald, 199n1, 199n5, 239

definitions, 6–7, 76, 79, 104–105, 116,  
117, 134. *See also* Hume, two  
definitions of cause

Della Rocca, Michael, 82n13, 83n31,  
83n35, 86n76

Descartes, René, 2–5, 24, 31–2, 33, 34,  
35–47, 48–64, 65–72, 81, 88,  
91, 93, 101, 102n4, 102n10;  
*Correspondence*, 34, 40, 41–42,  
46n17, 46n29, 55, 57, 58–59,  
61; *Le Monde*, 4, 36, 37, 43–44,  
52, 60; *Meditations*, 2, 35,  
49–51, 55–57, 68, 88; *Principles*  
*of Philosophy*, 35–44, 50, 51,  
55, 57, 58, 60, 62n6, 93, 102n4

determinism, causal, 7, 128, 154–157,  
201–203, 205–210

Digges, Thomas, 13

dispositions, 189–190

Donagan, Alan, 76



Dowe, Phil, 189, 199n4  
 Drake, Stillman, 19, 29n14, 29n25  
 Dryden, John, 176  
 Duhem, Pierre, 13  
 Dunn, John, 122n49  
 Durandus of Saint Pourçain, 31–35, 43, 44  
 duration, 2, 40–44, 50–51, 71, 76. *See also* time  
 dynamics, 74

## E

Eaton, William, 4, 5, 48–64  
 economy, principle of, 70, 87  
 Edgerton, Samuel, Jr., 29n30  
 error theory, causal. *See* projectivism  
 essences: 6–7, 42, 65–67, 76, 78–81, 90–91, 99–100, 109, 113–115, 174, 177, 254, 257; of body, 5, 40–41; of commonwealth, 107, 111, 115, 117; of God, 77, 93, 100; of man, 110; of mind, 5, 41–42  
 Eucharist, the, 97  
 Ewing, A. C., 217n19  
 explanation: causal, 1, 2, 15, 21–2, 109, 197, 202, 210, 212, 216, 223; mechanical, 1, 14–18, 51–52, 201–203, 205–209, 212–215; non-mechanical, 2, 17  
 expression, 76, 77–78, 94–95  
 extension, 3–4, 24, 32, 39–44, 48, 49, 51–52, 54–55, 59, 61, 68, 78, 91, 207

## F

Fair, David, 189  
 Fales, Evan, 200n7  
 Ferne, Henry, 120n10  
 Feyerabend, Paul, 13  
 Filmer, Sir Robert, 7, 107–112, 114, 115, 116  
 Flew, Anthony, 184n6  
 force, 2, 3, 16–18, 20, 32, 35, 50, 54, 57, 39–44, 90, 91–92, 93, 98–11, 127, 206, 207–208  
 Freddoso, Alfred, 31, 45n9  
 Friedman, Michael, 217n28, 218n41  
 Fritz, Kurt von, 119n10  
 Furet, François, 30

## G

Gabbey, Alan, 40–41, 86n72  
 Galileo, Galilei, 1, 13–30

Garber, Daniel, 39, 45n14, 48, 66, 70, 76, 82n12, 83n24, 83n28, 86n70, 102n9  
 Garrett, Don, 82n8, 82n18, 200n18  
 Gasking, Douglas, 199n2  
 Gaultier, David, 121n37  
 Geiger, Ido, 219n43  
 generation: 104, 114, 205, 208; of children, 108–110, 114; of commonwealth, 7, 111–112, 115, 116–118  
 Gilbert, William, 24, 28n12  
 Ginsborg, Hannah, 208, 218n32  
 Giovanni, Magini, 13  
 Glennan, Stuart, 199n3  
 God, 2–5, 31–8, 39, 43–44, 49–51, 52, 57, 59–61, 67–73, 76–79, 87–94, 97–101, 110, 220, 222, 252  
 Goodman, Nelson, 185n20  
 Grassi, Orazio, 29n20  
 Guerolt, Martial, 35, 45n14, 46–47n30  
 Guyer, Paul, 218n40, 219n43

## H

Hampton, Jean, 121n37  
 Hanson, Norwood, 189  
 Harris, Ian, 120n17, 120n24, 120nn26–27, 120nn29–30, 121n33  
 Harris, James, 162n35, 164n62  
 Haserot, Francis, 74  
 Hatfield, Gary, 46n28, 81n4  
 Hattab, Helen, 45n13, 48, 60–61, 63n16, 83n25  
 Heilbron, John, 28n1  
 Hennig, Boris, 8, 188–200  
 Heymans, Gerard, 194  
 Higgerson, Robert, 4, 5, 48–64  
 Hobbes, Thomas, 6–7, 8, 11, 104–122  
*horror vacui* 17–18, 20, 26  
 Horstmann, Rolf-Peter, 219n43  
 Hubner, K, 79  
 Huenemann, Charlie, 84n39  
 Hume, David, 7–9, 11, 27, 109, 123–165, 166–187, 188–191, 192–194, 196–199, 220, 221, 223–224, 225, 232–236, 237, 239, 242, 243–246, 249, 252, 255, 256–257; Copy Principle, 137, 140, 145, 155, 157, 170, 171, 172–174; Hume's Fork, 7, 177; missing shade of blue, 172–174; 'New Hume', 7–8, 123–126, 128, 132–134, 136,



137, 140–141, 142–143, 145, 146, 147, 150, 156, 157–158, 166–168, 182, 221; ‘Old Hume’, 7–8, 123–126, 128, 132, 134, 136, 139, 156–158, 167–168; two definitions of cause, 7–8, 123, 126, 127–128, 130–132, 136–137, 141, 149, 153–154, 168–169, 180, 182, 184; two definitions of necessity, 128, 130, 134–135, 153–156, 180, 231

Hunton, Philip, 120n10

Huygens, Christiaan, 24, 26

## I

imagination, 75–77, 115–118, 127, 133, 135, 145–146, 151, 170, 178

indeterminism, causal, 180, 186n47.  
*See also* determinism, causal; libertarianism

induction, 11, 24, 129, 135, 137, 139, 169, 177–178, 253–256

inertia, 5, 37, 70, 72, 73, 91, 226

intelligibility: nature of, 173, 174–176;  
 of mechanism, 4, 52–53, 61;  
 of mind-body relations, 5–6, 88–101; of necessary connexions, 7–8, 125, 132, 152, 154–155, 175, 180

interactionism, mind-body, 5–6, 87–90, 93, 101

interventionist notion of cause, 2, 24–26

## J

Jesseph, Douglas, 119n1, 119n4

Johnston, Samuel, 175–176

## K

Kail, Peter, 7, 124–126, 142–150, 151, 157, 158, 160n20, 160n25, 173–175, 177

Kant, Immanuel: 9, 191, 201–219, 255; *Critique of Judgement*, 201, 202, 204–207, 208, 209, 210, 211, 212, 214, 215; *Critique of Pure Reason*, 201, 202, 203–204, 206, 209, 210, 215; *Metaphysical Foundations of Natural Science*, 202–203, 207, 208–209, 210, 213, 215

Kavka, Gregory S., 121n37

Kemp Smith, Norman, 124, 151, 161n30, 240n6

Koestler, Arthur, 13

Koistinen, O, 84n42

Koyré, Alexandre, 84n52

Kraus, Jody S., 121n37

Kripke, Saul, 184n2

Kristeller, Paul Oskar, 82n10

Kuhn, Thomas, 26–27

## L

La Forge, Louis, 4, 48, 49, 51–54, 63n30

Lachterman, David, 66, 70, 80

Laerke, Mogens, 79, 83n36, 84n48, 85n53

Landes, M. W. 240n1

laws of nature, 9, 19, 31–32, 36, 38–44, 51, 60–61, 68, 69–70, 90–94, 100–101, 201–203, 204, 208, 209–214, 216, 223, 224, 227, 259n24; as arbitrary, 88–89, 91–94, 97, 98, 101, 259n24.

Le Pore, Ernest, 200n29

Leibniz, Gottfried Wilhelm, 5, 53, 74, 79, 83n30, 87–103

Leslie Affair, the, 1, 9, 220, 242

Levy, Ken, 45n14

Lewis, David, 188, 227

libertarianism, 134–137, 139, 158

Lilburne, John, 120n12

Lin, Martin, 82n8, 82n11

Livingston, Donald, 124

Locke, John, 94, 145, 164n61, 173, 174, 176, 259n24, 260n33

Loeb, Louis, 56

Löw, Reinhard, 216n4

## M

Machamer, Peter, 47n35, 64n35

Macherey, P., 83n36

Mackie, J. L., 178, 184n7, 185n20, 185n23, 196, 198, 260n28

Malebranche, Nicolas, 5, 49, 53, 54–55, 57–59, 61, 87–92, 93, 96, 101, 102n10, 169, 230

Mancosu, Paolo, 79

Marion, Jean-Luc, 45n10

Masham, Damaris, 6, 87, 88, 89, (97), (101)

maxims, 9, 201–203, 205–207, 208–209, 212–215

McFarland, John, 217n27

McLaughlin, Peter, 206–207, 208, 214, 216n5



meanings, 8, 125, 127, 132, 134, 136, 149, 168, 171, 174–176, 178.  
*See also* definitions; intelligibility  
 mechanics, 20–21, 22; Quantum, 98  
 mechanism, 1–2, 6, 8, 9, 11, 14–15, 17–18, 19, 24–26, 51–52, 61, 65, 86n70, 91, 201–203, 204–210, 212–215  
 Mellor, D. H., 188, 200n6  
 Melnick, Arthur, 218n41  
 metaphysics: and common life, 148; as first philosophy, 104–105  
 Mill, J. S., 260n28  
 Millican, Peter, 7–8, 123–165, 178, 180–181, 186n40, 186n45, 241n15  
 Mills, John, 221, 257nn1–2  
 mind-body causation, 4–5, 61, 87–104.  
*See also* body-body causation; substance, immaterial  
 miracles, 36, 37, 89–90, 92–94, 97–99, 101, 103n14, 252  
 modes, 39, 41–43, 51, 53, 54–59, 74–78, 81, 93, 99–100  
 monism, substance, 5, 65, 71–73, 99–100. *See also* Spinoza  
 Mont, Guidobaldo del, 13  
 More, Henry, 4, 55, 57, 58, 59–60, 61  
 Moscovici, Serge, 29n20  
 motion: 4–6 31–32, 35–44, 48–64, 69–70, 99, 207, 226, 247, 248; of falling bodies, 21, 23; as explanation of the tides, 14; and sensation, 89, 90–94, 95, 97, 98; quantity of, 35, 38–44, 71–72  
 Mutakallinûn, 32

## N

Nadler, 54, 55, 62n5, 62n15, 102n8, 185n18  
*natura naturans*, 72, 78, 102n7  
*natura naturata*, 72, 77–78, 102n7, 103n14  
 nature: course of, 11, 232, 236, 237, 251–252; Descartes's definition of, 43–4; material, 204–209, 213–216; state of, 7, 111; unity of, 203, 204, 209–212, 213–215. *See also*, essences; laws of nature; *natura naturans*; *natura naturata*  
 Naylor, R.H., 28n6  
 necessary connexions, 7–8, 11, 55, 90, 109, 112, 124–131, 133–138,

140–146, 149–151, 156–158, 167–170, 171, 173, 177–182, 192, 195, 221, 226, 233–234, 235, 244, 247, 252, 256–257.  
*See also* cause, secret or hidden; constant conjunctions; necessity, causal; powers  
 necessity: causal, 11, 55, 124, 126–132, 134–139, 141–143, 146–147, 149, 150–157, 166–170, 172, 177–181, 194–196, 203, 242, 244; conceptual, 196–198; mathematical, 254–257; metaphysical, 63n21, 100; moral, 137, 139  
 Nelson, J. O., 186n27  
 Newlands, Samuel, 75  
 Newton, Isaac, 19, 26, 74, 83n27, 83n30, 84n52, 183–184  
 nominalism, property, 226–228  
 Noonan, Harold, 168  
 Nyden, Tammy, 82n14

## O

occasionalism, 2–4, 5, 31–33, 39, 43, 48–64, 68, 87–8, 89, 90, 92–93  
 occult: causes, 24; forces, 2, 17, 39; qualities, 3, 145  
 organisms, 87, 89, 96–97, 98–99, 201, 204–205, 208, 214  
 Overton, Richard, 120n12

## P

Palmieri, Paolo, 28n8, 29n27  
 Pascal, Blaise, 17–18, 26–27  
 Penelhum, T., 187n55  
 perception: causal explanation of 5–6, 87–110; intellectual, 76–78; order of, 203–204; of causes, 10, 129–130, 138, 140, 143, 144, 151, 224, 234, 245, 245–246, 249, 256–7; of regularities, 236–237; *petites* or minute, 94–98  
 Perler, Dominik, 44n3  
 perservice, 38–39, 66, 70, 72, 73, 78.  
*See also* conatus; preservation  
 Pessin, Andrew, 45n13, 83n22  
 Peter, Joachim, 219n43  
 Peterson, James, 228  
 Phemister, Pauline, 4, 5–6, 87–103  
 pluralism, substance, 6, 99–101. *See also* Leibniz  
 Pollok, Konstantin, 217n28  
 Popkin, R. H., 187n56, 187n59



power: identity theory of, 10, 221, 224–230, 232–255; of God, 3, 5, 34, 36, 61, 68, 78; of material objects, 3, 9, 32, 34, 48–49, 55, 57, 58, 59–60, 90; of preservation, 2, 50; of sovereign, 107–108, 110, 116, 118; of words, 104; Reid's theory of, 9–10, 221, 222–224, 225; 'thick' ('hidden', 'secret', 'objective'), 125–127, 129, 132, 133–134, 135–136, 138, 140, 142–146, 148–149, 171–172, 174–75, 176, 178, 181–183, 221, 234–235; to persevere, 66, 71, 74

pre-established harmony, 5, 87–88, 89, 91–93, 96, 99

preservation: and creation, 2, 50; and the commonwealth, 7, 114–115; and giving of life, 114; of essence, 5, 65, 69–70, 72, 73 (*see also* conatus); of quantity of motion, 57, 59, 69, 71–72

primary qualities. *See* extension; motion; secondary qualities

process: causal, 8, 20, 188–190, 192–4, 198–199, 204, 210–211; latent, 79–80; union-forming, 244

projectivism, causal, 145, 150, 162n40, 164n54, 168, 170

Prynne, William, 120n12

Psillos, Stathis, 9–10, 220–241, 257n1

Ptolemaic astronomy, 14, 15, 22

## Q

Quarfood, Marcel, 219n52

## R

Radner, Daisie, 56

Raijva, Suma, 219n43

Rang, Bernhard, 216n10

Read, Rupert, 185n20

realism: anti-, 142, 166–167, 170, 177, 180–181; empirical, 7–8, 181; quasi-, 177, 181; sceptical, 7–8, 124–125, 133–137, 141–143, 146–148, 157, 166–167, 169–170, 177, 181, 221. *See also* nominalism, property

re-creation. *See* cinematic view of Divine conservation; creation

Redondi, Pietro, 29n20

Reeves, Eileen, 28n11

reflective judgement, 9, 202, 211–215

Regis, Pierre-Sylvain, 61

regularity theory of causation, 9–10, 11, 123–125, 153, 166–167, 180, 221–222, 223–225, 227, 228–231, 236–239, 242

Reid, Thomas, 9–10, 164n61, 220, 221, 222–224, 225, 231, 232, 238–239, 246, 257

Renn, Jürgen, 28n5

rest, 45n12, 50, 53, 57, 58–59, 70, 79, 84n47. *See also* motion

Revius, Jacobus, 38

Robinson, J. A., 185n10

Rödl, Sebastian, 200n25

Rohault, Jacques, 61

Ross, Alexander, 110

Rosso, Lucio, 29n28

Roth, A. S., 185n20

Russell, Bertrand, 195, 259n21

Russell, Paul, 169

## S

Sagredo, Giovanni, 16, 18

Salmon, Wesley, 189

Sandis, Constantine, 7–8, 166–187

Sans, Georg, 217n10

Schelling, Friedrich, 85n53

Schlick, Moritz, 196

Schliesser, Eric, 4, 65–86

Schmaltz, Tad, 2–4, 31–47, 48, 50–51, 56, 64n35, 82n22

Scholasticism. *See* Aristotelianism

Scott, Jonathan, 120n12, 120n15

Scruton, Roger, 184n9

secondary qualities, 2, 24, 26, 95–96, 246

separability: of cause and effect, 8, 145, 150–151, 189, 190–191, 197, 198–199; conceptual, as a guide to real, 68–70

Shabel, Lisa, 261n43

Shea, William R., 29n22

Shein, Noa, 73–74, 84n38, 84n45, 84n53, 85nn57–58, 85n62, 85n64

Shepherd, Lady Mary, 1, 9, 10–11, 241n16, 242–261

Simmons, Alison 82n22

Skinner, Quentin, 107, 120nn12–13, 120n16, 121n41, 121n45

Sleigh, Robert, C., Jr., 12n2, 62n4

Smith, Sir Thomas, 106–107, 112, 113

Sorell, Tom, 119n2

sovereignty, 7, 107–111, 115–118



Spinoza, Baruch, 5, 6, 65–86, 93, 94, 99–101

Spragens, Thomas A. Jn., 119n2

Stanton, Timothy, 6, 104–122

Stevin, Simon, 26

Stewart, Dugald, 220, 255–256, 257

Stillingfleet, Edward, 110

Stoneham, Tom, 186n36

straitjacket, 172, 252

Strawson, Galen, 124–125, 142, 148, 160n20, 161n30, 163n42, 163n44, 174, 177, 240n5, 240n6

Strawson, P. F. 187n54

Stroud, Barry, 184n5, 234–235

Sturm, Johann Christoph, 53

Suárez, Francisco, 31–35, 36, 37, 39–40, 43, 44

substance: 7, 11, 67–68, 72–73, 74–76, 97–101, 226–228, 247; created, 2–3, 31–35, 59–61, 68–69, 71–72, 90; infinite (God), 2, 68–9, 72, 222; immaterial (mental), 2, 5–6, 41–42, 59–60, 68, 70–72, 87–84, 100–1; material (body), 2–5, 31–32, 35, 37–42, 43–44, 48–49, 50–52, 54–55, 57–61, 60–61, 68–70, 75, 87–95, 100–101, 109–11, 128–132, 149–150, 153, 154, 207–208, 213, 227

sufficient reason, principle of, 70, 71, 250–251

## T

Tarlton, Charles D., 120n13

Taylor, Richard, 189

teleology, 11, 82n17, 84n48, 205, 214–215

thick connexions. *See* necessary connexions

tides, explanation of, 2, 14–15, 26

time, 35, 50–51, 68–72, 74, 204, 206.

*See also* duration

Torricelli, Evangelista, 17–18, 26

Tuck, Richard, 119n8

## V

Vampoulis, 70, 73

Vaughan, G. 122n47

Viljanen, Valterri, 73–74, 79, 82n8, 84n41

von Wright, G. H., 184n8

## W

Waard, Cornelis de, 29n22

Watkins, Eric, 218n33

Welsh, David, 228, 231, 239, 240n8, 257n2

Whewell, Robert, 258n5

Williams, W. H., 186n25

Wilson, Catherine, 69

Winkler, Kenneth, 124–126, 168

Wittgenstein, L. 181

Wolfson, 74, 76, 82n9

Wolsterstorff, Nicholas, 240n7

Woolhouse, Roger, 184n3

Wootton, David, 1–2, 13–30

words: as causes, 104, 117; and images of, 115–6, 117, 118; categories of, 115–116

Wrede, W., 187n57

Wright, John, P., 7, 30n39, 124–125, 142, 150–157, 158, 160n20, 163n44, 163n49, 240n5

## Y

Yolton, John, 124, 155, 160n26

## Z

Zammito, John, 217n27

Zarka, Yves Charles, 119n4

Zukert, Rachel, 218n40